Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1964

The sensitivity of optimal strategies in finite statistical decision problems to variations in parameters.

Bovey, R. L.

The Johns Hopkins University

http://hdl.handle.net/10945/13014

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

NPS ARCHIVE
1964
BOVEY, R.

Library
U. S. Naval Postgraduate School
Monterey, California

THE SENSITIVITY OF OPTIMAL STRATEGIES IN FINITE SEQUENTIAL DECISION PROCESSES TO VARIATIONS IN PARAMETERS

By

R. L. Bovey

A dissertation submitted to the Advisory Board
of the School of Engineering of The Johns Hopkins
University in conformity with the requirements
for the degree of Doctor of Philosophy

Baltimore
The Johns Hopkins University
1964

NPS ARCHIVE
1964
BOVEY, R

Library
U. S. Naval Postgraduate School
Monterey, California

Abstract

The paper is devoted to statistical decision problems involving a finite number of states of nature, experimental indications, and terminal actions. A device, called a "criterion matrix," for the specification of the optimal (Bayes) strategy for such a problem is proposed. Then methods are presented for analyzing the sensitivity of this strategy to variations in parameters of the problem which influence the criterion matrix linearly and which are subject to linear constraints.

Value estimation techniques based on the von Neumann-Morgenstern theory of utility using a decision maker's choices among lotteries and techniques presented by Churchman et al. in their Introduction to Operations Research are combined to illustrate that the nature of the imprecision often associated with value estimates is such that sensitivity analysis of the sort described can be employed to determine if the estimates available in a given situation are adequate to single out one strategy as being optimal.

Sensitivity analysis is also an adjunct to informal analysis in that it can be used to specify the values of parameters for which particular strategies are optimal when these parameters influence the criterion matrix linearly. This approach is useful in determining if imprecise knowledge of one or more of the probability distributions involved in

the decision problem is sufficient for an unambiguous specification of an optimal strategy.

Special methods of sensitivity analysis for variations in a single parameter are described, and in conjunction with sensitivity analysis for variations in several parameters, general methods are presented for investigating a convex set defined by linear constraints. Included are methods for revealing inconsistencies in the system of constraints, redundancy of individual constraints, and unboundedness of the set as well as a discussion of the specification of a set of points which is sufficient to define the convex set.

Acknowledgments

The author wishes to express his appreciation to his referees, Dr. Charles D. Flagle, Professor of Operations Research and Industrial Engineering, and Dr. Carl F. Christ, Professor of Political Economy for their guidance and constructive criticism in the preparation of this dissertation. The author is indebted to Mr. William H. Walton for the discussions which led to the ideas on minimax error strategies in classification problems presented in Chapters 4 and 5. Finally the author wishes to express his appreciation to Mrs. Mary Kilburn for her assistance in preparing the manuscript.

This study was carried out while the author was assigned to The Johns Hopkins University under the Junior Line Officer Advanced Scientific Educational (Burke) Program of the United States Navy, and was partially supported by U.S. Public Health Service Grant Number HM 00279-01.

Table of Contents

Page

Chapter 1: Finite Statistical Decision Theory 1

I. Definitions 3

II. Method of Solution 9

Chapter 2: Value Estimation Techniques 17

I. Definitions 18

II. Simplifications in the Statement of Loss Matrix Entries 21

III. Some Methods for Obtaining Loss Matrix Entries. . 25

A. Obtaining Loss Matrix Entries When the Value Function is a Sum 26

B. Obtaining Loss Matrix Entries When the Value Function is not a Sum 44

IV. Summary 49

A. Assumptions of Sensitivity Analysis Methods . 49

B. Review of Value Estimation Methods 50

Chapter 3: Sensitivity Analysis for Variations in Several Parameters 58

I. Investigation of Strategies Optimal for Given Constraints on Parameters 60

A. General Discussion 60

B. Example of Sensitivity Analysis 70

II. Specification of Optimality Region for a Given Strategy 85

A. General Discussion 86

B. Example of Sensitivity Analysis 88

Page

Chapter 4: Sensitivity Analysis for Variations in a Single Parameter 94

I. Analysis When D is Arbitrary 96

A. Preliminary Discussion 96

B. Analysis Procedure 103

C. Examples of Sensitivity Analysis 115

II. Analysis When D Contains Non-zero Elements in Only One Row 127

A. Analysis of Changes in the Optimal Strategy . 128

B. The Minimum Expected Loss 133

C. Example of Sensitivity Analysis 138

Chapter 5: Sensitivity Analysis Applied to Probability Distributions 147

I. The Criterion Matrix as a Linear Function of the Elements of a Probability Distribution . . . 147

A. A Prior Distribution over States of Nature and Conditional Distributions over Indications 148

B. A Prior Distribution over Indications and Posterior Distributions over States of Nature 149

C. Applicability of Sensitivity Analysis to Probability Distribution Parameters 151

II. Examples of Sensitivity Analysis 152

A. Sensitivity Analysis for Elements of a Prior Distribution over States of Nature . . . 154

B. Sensitivity Analysis for Elements of Conditional Distributions over Indications . 162

Page

Appendix: Development of a Method for the Investigation of a Convex Set Defined by Linear Constraints . 172

I. Reduction of a System of Both Equalities and Inequalities to One Involving Only Inequalities . . 173

A. Description of Constraints 173

B. Reduction Operation 174

II. Basic Theorems 176

A. Theorems on the Consistency of Systems of Inequalities 176

B. Theorem on Unboundedness 182

C. Theorems on Redundancy of Constraints 183

III. The Use of Linear Programming Techniques to Obtain a First Extreme Point Solution, Check Consistency, and Detect Unboundedness 189

A. The Linear Programming Technique 189

B. Results of an Application of the Simplex Algorithm 196

C. A Necessary and Sufficient Condition for Unboundedness 210

IV. Testing Constraints for Redundancy 215

A. General Method for the Redundancy Test 216

B. Application of the Dual Simplex Algorithm . . . 219

C. Other Criteria for Redundancy and Nonredundancy of Constraints 220

V. Generating Extreme Points 222

A. Recording Extreme Points 222

Page
B. Determining Adjacent Extreme Points 224
C. Procedure 229
VI. Incorporation of New Constraints 232
VII. Example 235

Chapter 1

Finite Statistical Decision Theory

In the past decade interest in statistical decision theory has grown among statisticians and others interested in the quantification of decision problems in such varied fields as military tactics, economics, and medicine. In a rough way we can describe the statistical decision problem as the problem of deciding what to do on the basis of incomplete evidence. In non-statistical decision situations where our evidence is complete, we can simply select the action which leads to the most desirable of the attainable outcomes, but in statistical decision situations our inability to observe the world perfectly prevents us from predicting with certainty the outcome which will result from our actions. Thus we must consider the values of all possible outcomes and the probability of achieving each outcome when we select a particular action. The problems with which the theory of statistical decisions deals are notable principally in that the notions of "value" and "probability" of an outcome can be suitably quantified.

The statistical problems associated with the theory (primarily the estimation of properties of probability distributions) have been studied intensively and profitably although many unsolved problems remain. However efforts to quantify the concept of a "good" or "bad" act have been less productive,

chiefly because of difficulties inherent in the intangible and subjective entities called "values," "utilities," or "costs." In environments where only one objective measure of value (e.g. money) is relevant, it is sometimes possible to quantify precisely. However precise quantifications are all but nonexistent in environments where more than one measure of value is relevant or where subjective values must be considered.

The most common applications of the theory to date in medicine have been to problems of classifying patients as belonging in one or another diagnostic category on the basis of various test results. At least one reason for the heavy concentration of applications in this particular area is that an act of classification can be regarded as "correct" or "erroneous." With such a dichotomy the problem of quantifying the efficacy of an act becomes trivial (see example 4 in Chapter 4), and the classification scheme can be based on firmly established statistical principles.

Flagle and Lechat (1963) have presented a case for the quantitative consideration of public health problems in which a simple dichotomy is patently inadequate and have made the point that even an imprecise quantification of the "goodness" or "badness" of acts will sometimes be adequate for the purpose of specifying a "best" decision. The purpose of this

paper is the development of methods to determine if the quantification made in a specific problem is precise enough to single out one decision rule as being best or to provide information about those rules which could conceivably be best within the framework of the available quantification. This will be done through the device of analyzing the sensitivity of an optimal solution of a statistical decision problem to variations in certain parameters of the problem.

We make no pretense of solving all such problems for certain restrictions have been made on the class of problems considered, the first of which is confining our attention to "finite" statistical decision theory. Our immediate goal is to outline this theory briefly and to present a solution procedure yielding results in a form amenable to a posterior analysis of the sensitivity of the solution to variations in certain parameters of the problem.

I. Definitions: We will describe a finite statistical decision problem in terms of seven basic components, "states of nature," "experimental indications," "terminal actions," a "weighting function," a "conditional probability distribution," a "prior probability distribution," and "strategies."

The first component is the set $\{\theta_n\}$ of a finite number of "states of nature" which we assume to be mutually exclusive and exhaustive in the sense that one and only one of θ_1, ...,

$\theta_n, \ldots, \theta_N$ can occur. The states of nature are presumed to incorporate those aspects of the outside world which are relevant to the problem at hand. The example to be used here concerns a decision to be made in the course of treating leprosy patients and was presented by Flagle and Lechat (1963). In it the only relevant fact about the outside world is whether or not a patient with manifest symptoms of leprosy is contagious. Thus we have two states of nature, θ_1 being the state of a contagious patient and θ_2 being the state of a noncontagious patient.

The second component of the problem is the set of "experimental indications" of the state of nature which we will denote by $\{x_i\}$ or $x_1, \ldots, x_i, \ldots, x_I$. As with states of nature, the indications are assumed to be finite in number and to be mutually exclusive and exhaustive. In our example the diagnostic "experiment" involves the microscopic examination of smears and can produce either a positive (contagious) indication, x_1, or a negative indication, x_2.

The third component is the set of "terminal actions" which we will denote by $\{a_j\}$ or $a_1, \ldots, a_j, \ldots, a_J$. Again these actions are assumed to be finite in number and to be mutually exclusive and exhaustive. In particular the set will often include a "null act" which is to say that nothing is done in response to the experimental results. In our example three

actions are considered. The first, a_1, is to hospitalize the patient, the second, a_2, is to treat the patient as an outpatient, and the third, a_3, is to discharge the patient without further treatment.

The fourth component of the problem is a "weighting function" which in a finite problem can be presented as a "loss matrix" or "utility matrix" depending upon one's viewpoint. We will adopt the terminology of losses because we have in mind the application of this theory to public health problems where the very existence of disease represents a costly departure from the norm. We denote the loss matrix by L. It is a J x N matrix (i.e. J rows and N columns), and the entry in row j and column n l_{jn} is taken to be a quantitative representation of the loss suffered when action a_j is taken in the presence of state of nature θ_n. For future reference we designate the n-th column of L by l_n which is a J-dimensional column vector containing the losses associated with state of nature θ_n. We defer a discussion of the methods of obtaining the loss matrix to the next chapter and simply take the following loss matrix as given for our example.

$$\begin{array}{rc}
 & \text{States of Nature} \\
 & \begin{matrix}\theta_1 & \qquad & \theta_2\end{matrix} \\
\text{Actions}\quad \begin{matrix}a_1\\ a_2\\ a_3\end{matrix} & \begin{pmatrix}1.0 & \qquad & 1.0\\ 1.2 & \qquad & .4\\ 3.0 & \qquad & .5\end{pmatrix}
\end{array}$$

The fifth component is a "conditional probability distribution" which in the finite problem can be presented as an $N \times I$ matrix P where the ni-th element of P is p_{ni} = $\text{Prob}(x_i | \theta_n)$, that is the probability that the experiment will yield indication x_i when θ_n is the true state of nature. Being probabilities the elements of P are non-negative, and to formalize the statement that given the state of nature one of the indications must be produced by the experiment we write

$$P_n' 1_I = 1 \qquad \text{for } n = 1, \ldots, N$$

where 1_I is an I-dimensional column vector each element of which is unity and P_n' is the n-th row of P. Here we introduce the convention of denoting row vectors by the transpose symbol. In our example we will take as our conditional probability the distribution the following matrix.

$$\begin{array}{ccc} & & \text{Indications} \\ & & \begin{array}{cc} x_1 & x_2 \end{array} \\ \text{States of Nature} & \begin{array}{c} \theta_1 \\ \theta_2 \end{array} & \left(\begin{array}{cc} .8 & .2 \\ 0 & 1.0 \end{array} \right) \end{array}$$

The sixth basic component of the problem is a "prior distribution" which can be presented here as an N-dimensional row vector Q' the n-th element of which q_n gives the prior ("prior" to the conduct of the experiment) probability that

θ_n is the true state of nature. Each element of Q' is non-negative and $Q'1_N = 1$ since one state of nature must obtain. In our example we will take as the prior distribution the following vector,

States of Nature

$$\begin{array}{cc} \theta_1 & \theta_2 \\ (\,.1 & .9\,) \end{array}$$

Instead of a prior distribution over states of nature and a distribution over experimental indications conditional on the true state of nature, a problem may be encountered in which it is more convenient to obtain a distribution over experimental indications and a "posterior" distribution over the states of nature, i.e. a distribution which represents the probabilities that θ_n $(n = 1, \ldots, N)$ is the true state of nature given that indication x_i $(i = 1, \ldots, I)$ was observed. This alteration presents no particular difficulties and a modification of the solution technique suitable to this situation will be presented.

The seventh and last basic component of the problem is the set of strategies. In general statistical decision theory both "pure" and "mixed" strategies are considered, but here we consider only the former and dispense with the adjective "pure." By $\{s_k\}$ or $s_1, \ldots, s_k, \ldots, s_K$ we denote the set of strategies, each of which is a rule stating the terminal action to be taken when each indication is observed. Formally

a strategy is a "many-to-one" mapping from the set $\{x_i\}$ to the set $\{a_j\}$ and can be represented by an I x J "strategy matrix" S_k in which the entries s_{ij} have the following properties.

1. $s_{ij} = 1$ means "if x_i is observed take a_j."
2. $s_{ij} = 0$ means "if x_i is observed do not take a_j."
3. $_{j}\, s_{ij} = 1$, i.e. there is but a single one in each row of S_k.

In some cases it is more convenient to represent s_k by an I-tuple, the i-th element of which is the j for which $s_{ij} = 1$ in S_k.

As an illustration we may consider a strategy (let us call it s_6) in our example having the following strategy matrix.

$$\begin{array}{cc} & \text{Actions} \\ & \begin{array}{cccc} & a_1 & a_2 & a_3 \\ \text{Indications} \quad x_1 & 0 & 1 & 0 \\ \phantom{\text{Indications}} \quad x_2 & 0 & 0 & 1 \end{array} \end{array}$$

In words this strategy is to treat as an outpatient a patient for whom a positive indication is obtained and to discharge a patient for whom a negative indication is obtained. There are nine possible strategies for our example as shown by the pairs below.

s_1: (1, 1)
s_2: (1, 2)
s_3: (1, 3)
s_4: (2, 1)
s_5: (2, 2)
s_6: (2, 3)
s_7: (3, 1)
s_8: (3, 2)
s_9: (3, 3)

In general the number of strategies will be given by

$$K = (J)^I$$

since each of the elements of the I-tuple can take any one of J values. Because the number of strategies is quite large even in small problems, it is incumbent upon us to arrange our calculations in as efficient a way as possible. We now turn to the formulation of a calculation procedure.

II. Method of Solution: If we decide to use a particular strategy having the matrix S_k and if θ_n is the true state of nature, we can construct an I-dimensional column vector T_{kn}, the i-th element of which gives us the loss incurred if x_i is observed. We obtain T_{kn} by

$$T_{kn} = S_k L_n$$

We note for future reference that if we augment T_{k2} on the right of T_{k1} etc. we form an I x N matrix T_k which can be written as

$$T_k = (T_{k1}, T_{k2}, \ldots, T_{kn}, \ldots, T_{kN}) = S_k L$$

In our example the vector T_{61} is given by

$$T_{61} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1.0 \\ 1.2 \\ 3.0 \end{pmatrix} = \begin{pmatrix} 1.2 \\ 3.0 \end{pmatrix}$$

and the matrix T_6 is given by

$$T_6 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1.0 & 1.0 \\ 1.2 & .4 \\ 3.0 & .5 \end{pmatrix} = \begin{pmatrix} 1.2 & .4 \\ 3.0 & .5 \end{pmatrix}$$

Furthermore if we use a_k the expected loss conditional on θ_n being the true state of nature is given by

$$y_{kn} = P'_n T_{kn} = P'_n S_k L_n$$

and we can collect the y_{kn} for all n into an N-dimensional column vector Y_k. In our example we have

$$y_{61} = (.8 \quad .2) \begin{pmatrix} 1.2 \\ 3.0 \end{pmatrix} = 1.56$$

$$Y_6 = \begin{pmatrix} y_{61} \\ y_{62} \end{pmatrix} = \begin{pmatrix} 1.56 \\ .5 \end{pmatrix}$$

To determine the "unconditional" expected loss of the k-th strategy $E(s_k)$ we use the prior distribution Q' and compute

$$E(s_k) = Q'Y_k$$

In our example we have

$$E(s_6) = (.1 \quad .9) \begin{pmatrix} 1.56 \\ .5 \end{pmatrix} = .606$$

The optimal or "Bayes" strategy is defined to be that strategy for which the expected loss is minimized. Formally an optimal strategy s_o is a strategy such that

$$E(s_o) = \underset{k}{\text{Min}}\, E(s_k)$$

Had we chosen to work with utilities instead of losses, the optimal strategy would be the one that maximizes expected utility. As noted previously the choice of approach is arbitrary. The statements in the remainder of this chapter will remain valid if we replace "Min" by "Max" throughout.

An alternate statement of the optimal strategy may be developed as follows. Let us define $\not{t}$ to be a square matrix having the elements of the prior distribution arranged along its principal diagonal (i.e. upper left to lower right) while all off-diagonal elements are zero. We may then perform the matrix multiplications

$$\Phi PT_k = \Phi PS_k L = \begin{pmatrix} q_1 y_{k1} & \cdots & q_1 P'_1 T_{kn} & \cdots & q_1 P'_1 T_{kN} \\ \vdots & & \vdots & & \vdots \\ q_n P'_n T_{k1} & \cdots & q_n y_{kn} & \cdots & q_n P'_n T_{kN} \\ \vdots & & \vdots & & \vdots \\ q_N P'_N T_{k1} & \cdots & q_N P'_N T_{kn} & \cdots & q_N y_{kN} \end{pmatrix}$$

Upon taking the trace of the resulting matrix we find that

$$\mathrm{tr}(\Phi PS_k L) = \sum_{n=1}^{n=N} q_n y_{kn} = Q'Y_k = R(s_k)$$

Therefore we can specify the optimal strategy by its strategy matrix S_o where

$$\mathrm{tr}(\Phi PS_o L) = \underset{k}{\mathrm{Min}}\ \mathrm{tr}(\Phi PS_k L)$$

The theorem stating, "The trace of the product of matrices is invariant under a cyclical permutation of the matrices," is proved in Graybill (1961) on page 7, and with it we can restate the criterion for the optimal strategy matrix as

$$\mathrm{tr}(S_o L\Phi P) = \underset{k}{\mathrm{Min}}\ \mathrm{tr}(S_k L\Phi P)$$

which suggests the following computational procedure for determining the optimal strategy in a finite statistical decision theory problem.

1. Compute the J x I "criterion" matrix, $C = L\Phi P$.

The element in the j-th row and i-th column of C is

$$c_{ji} = \sum_{n=1}^{n=N} l_{jn} d_n p_{ni}$$

2. Circle $c_i = \underset{j}{\mathrm{Min}}\ c_{ji}$ for each column of C.

3. Construct the transpose of the optimal strategy matrix S'_o by placing ones in the cells of a J x I matrix that correspond to the circled elements of C and placing zeros in the other cells of this matrix.

4. The minimum expected or "Bayes" loss is then given by

$$\mathrm{tr}(S_o C)$$

where S_o is obtained by transposing the matrix of step 3. In some cases the Bayes loss may be computed more readily by noting the value of c_i at the foot of each column in the course of step 2 and then summing over all i.

The matrix S_o obtained by this procedure is the optimal strategy matrix because when we multiply the i-th row of S_o and the i-th column of C, we "pick out" the minimum element of that column and enter it on the principal diagonal (row i and column i) of the matrix S_oC. Therefore when the trace operation sums the diagonal elements of this matrix, it produces $\sum_{i=1}^{i=I} c_i$. Any matrix, S_k, different from S_o would not

"pick out" the minimum element of one or more columns of C and consequently $tr(S_o C) \leq tr(S_k C)$ for all $k = 1, \ldots, K$. If for some column i of C, c_i is not unique there will be alternative optimal strategies.

As an illustration of the procedure let us compute the optimal strategy for the examples of this chapter.

1. Compute the criterion matrix.

$$C = L\phi P = \begin{pmatrix} 1.0 & 1.0 \\ 1.2 & .4 \\ 3.0 & .5 \end{pmatrix} \begin{pmatrix} .1 & 0 \\ 0 & .9 \end{pmatrix} \begin{pmatrix} .8 & .2 \\ 0 & 1.0 \end{pmatrix} = \begin{pmatrix} .08 & .92 \\ .096 & .384 \\ .24 & .51 \end{pmatrix}$$

2. Circle $c_i = \underset{j}{\text{Min}}\, c_{ji}$ for $i = 1,2$.

$$\begin{pmatrix} \textcircled{.08} & .92 \\ .096 & \textcircled{.384} \\ .24 & .51 \end{pmatrix}$$

3. Construct S'_o.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ which is } S'_2$$

4. The minimum expected loss is

$$tr \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} .08 & .92 \\ .096 & .384 \\ .24 & .51 \end{pmatrix} = tr \begin{pmatrix} .08 & .92 \\ .096 & .384 \end{pmatrix} = .464$$

The particular example we have used is also an illustration of a phenomenon that can be used to reduce computations in some problems. Returning to the loss matrix we note that action a_3 "dominates" action a_2 in the sense that the loss incurred by taking a_2 is less than the loss incurred if a_3 were taken for every state of nature. Thus we can conclude that it will never be reasonable to use a_3. In general we can say that given a_j and $a_{j'}$ such that $l_{jn} \leqq l_{j'n}$ for $n = 1, \ldots, N$, $a_{j'}$ dominates a_j and $a_{j'}$ may be discarded.

We turn now to consider the alterations in our method which will be required when we have a distribution over the indications and a posterior distribution over the states of nature. We note first that the N x I matrix ϕP gives the joint probability distribution of $\{x_i\}$ and $\{\theta_n\}$, i.e. the element in the n-th row and i-th column of this matrix is Prob (θ_n, x_i). In the previous case the joint probability was given by

$$\text{Prob } (\theta_n, x_i) = \text{Prob } (x_i | \theta_n) \text{ Prob } (\theta_n)$$

while in this case the joint probability is given by

$$\text{Prob } (\theta_n, x_i) = \text{Prob } (\theta_n | x_i) \text{ Prob } (x_i)$$

Therefore to obtain the matrix of the joint distribution we perform the matrix multiplication $Q\phi$ where now Q is an N x I matrix the n-th element of which is Prob $(\theta_n | x_i)$ and ϕ is a

diagonal matrix having the distribution Prob (x_i) on its principal diagonal.

Thus to cope with this alternate form of our information on the probabilities, we need only alter step 1 of the procedure to read, "Compute the J x I matrix $C = LQ\mathcal{P}$." The remaining steps can be carried out as before.

We note that in this case

$$c_{ji} = \sum_{n=1}^{N} l_{jn} \text{Prob}(\theta_n | x_i)\ \text{Prob}(x_i) = \text{Prob}(x_i) \sum_{n=1}^{N} l_{jn} \text{Prob}(\theta_n | x_i)$$

The selection of c_i for each column of C will be independent of Prob(x_i). Therefore if our purpose is only to determine δ_c, we may use LQ as our matrix C. However to use the procedure to determine the minimum expected loss, we must use $LQ\mathcal{P}$.

This discussion by no means covers the whole of statistical decision theory. However it fulfills our need for a compact summarization of the data that will be relevant to our sensitivity analysis. For a general discussion of statistical decision theory we refer to Raiffa and Schlaifer (1961) who also offer other solution techniques which may be better adapted to specific problems than the method presented here. Chernoff and Moses (1959) present a more elementary exposition of the theory and describe a graphical solution procedure for the case N=2 which has great intuitive appeal because it exposes some of the important relations among various components of the problem.

Chapter 2

Value Estimation Techniques

At the outset we note that either the term "value" with its connotation of desirability or "cost" with its connotation of undesirability may be used. We will employ the former in the general discussion to conform to the usage of the authors whose ideas we discuss and the latter in our examples to emphasize the interchangeability of the terminology.

In this chapter we will discuss only a few of the techniques which have been suggested for the quantitative determination of values. Since the efficacy of any formally consistent technique is determined by the environment in which it is to be applied, the inclusion of particular techniques here does not imply that they are intrinsically better than some other technique. All of the techniques discussed here contain a subjective element of preference because they are based on choices between alternatives made by an individual or group. In some cases (usually those in which only monetary considerations are relevant) accounting procedures can provide all requisite value information, and techniques such as those discussed here need not be employed.

Our primary purpose in discussing techniques is to show that they generate value estimates in a form suggesting the

Chapter 2

Value Measurement Techniques

At the outset we note that either the term "value" with its connotation of desirability or "want" with its connotation of unconditionality may be used. We will employ the former in its general connotation to conform to the usage of the authors whose ideas we discuss and the reader is not expected to appreciate the interchangeability of the terminology.

In this chapter we will discuss only a few of the many techniques that have been suggested for the quantification of human scales of values. Since the existence of any formally constructed techniques is determined by the environment in which it is to be applied, the inclusion of particular techniques here does not imply that they are intrinsically better than some other techniques. All of the techniques discussed here contain a fundamental element of preference between that are based on choices between alternatives made by an individual or group. In each case (usually those in which only ordinary considerations are involved) measuring preferences may provide all pragmatic value information and techniques such as these discussed here need not be employed.

Our primary purpose in discussing techniques is to show that they generate value estimates in a form compatible with the desirability of a certain type of sensitivity analysis. Whether or not another technique will yield estimates in a similar format can only be determined by examining that technique.

At present a considerable argument flourishes concerning the philosophical foundations of value theory. Some stand must be taken on the question of what is "value" or "cost" else no presentation of estimation techniques is possible. The discussion here is based primarily on the ideas of C. W. Churchman and R. L. Ackoff (1957) and (1962), although it purposely will remain vague concerning philosophical questions. The point to be emphasized is not agreement or disagreement with these particular concepts of value, but is whether or not the concepts and techniques we adopt will lead to a set of value estimations which can be summarized formally in a way that makes them amenable to sensitivity analysis by the methods to be proposed in subsequent chapters.

I. Definitions: As a point of departure we assume that we are interested in the value structure associated with an individual or group called the "decision maker," that we can summarize the desires and apprehensions of the decision maker with regard to the problem at hand in a list of "objectives," and furthermore that these objectives admit to quantification.

Thus in the leprosy example the decision maker may be the individual or committee with the responsibility for public

health administration in a region. His or their objectives relative to an individual patient might be:

1. Minimize the number of new infections arising from a contagious individual.

2. Bring the disease under control in the patient.

3. Minimize the time the patient is required to be away from his productive labors for treatment.

4. Minimize the monetary cost of treatment in order to free funds for attacks on other public health problems.

In such real situations the identification of the decision makers and the specification of objectives are more art than science and can be difficult tasks. Bearing this in mind we will simply state formally that the decision maker has M objectives $O_1,\ldots,O_m,\ldots,O_M$. Each objective is attainable in varying degrees. This degree of attainment will be called, following Churchman et al. (1957), the "measure of efficiency for O_m" and is measured as a real number z_m. Three special cases are of interest in that they dictate the treatment of the problem.

1. Continuous quantitative objectives for which $\text{Min}(z_m) \leq z_m \leq \text{Max}(z_m)$ where usually $\text{Min}(z_m)=0$ and sometimes $\text{Max}(z_m)$ can be arbitrarily large. In the example above the third objective would be of this type. For convenience we denote this set by $O_1,\ldots,O_{M_1}$.

2. Discrete quantitative objectives for which $\text{Min}(z_m) \le z_m \le \text{Max}(z_m)$ where usually $\text{Min}(z_m)=0$ so that $z_m=0, 1, 2, \dots, \text{Max}(z_m)$, and sometimes $\text{Max}(z_m)$ can be arbitrarily large. In the example above the first objective would be of this type. We denote this set of objectives by $O_{M_1+1}, \dots, O_{M_2}$.

3. Qualitative objectives for which $z_m = \text{Min}(z_m)$ or $\text{Max}(z_m)$ where usually $\text{Min}(z_m)=0$ signifies that the objective is not fulfilled while $\text{Max}(z_m)=1$ signifies that it is. The second objective above would be of this type. We denote this set by $O_{M_2+1}, \dots, O_M$.

We assume that the list of objectives is rich enough so that we may describe all that is relevant about the outcome of an action a_j in the presence of θ_n by an M-tuple $Z_t = (z_{1t}, \dots, z_{Mt})$. The set of mutually exclusive and exhaustive outcomes, Z, is the set of M-tuples obtained by allowing each z_m to take on all possible values. We assume that at some time following the application of a_j in the presence of θ_n, it is possible to determine specific values for the M measures of efficiency, that these results are sufficient to determine the value or cost of this particular application of a_j under θ_n, and that if two outcomes can be described by the same M-tuple they are equivalent in the mind of the decision maker regardless of what other factors might be used to differentiate them. Here

we have made a distinction between "outcome" and "objective" which is not usual although it will be useful in the remainder of this chapter. The definitions of the two concepts are circular in that we may define an outcome as the M-tuple that gives specific measures of efficiency relative to the objectives, or we may define the objectives as being constructs such that their measures of efficiency are necessary and sufficient to differentiate among outcomes.

The entry l_{jn} in the loss matrix is defined as a "measure of effectiveness" of a_j under θ_n. Therefore in very general terms we can conceive of this entry as a function of the action, the state of nature, the values of the outcomes, and the probabilities of the various outcomes. Symbolically we say,

$$l_{jn} = f(u(Z),\ F(Z),\ a_j,\ \theta_n)$$

where u(Z) is a value function defined on the set Z, and F(Z) is a joint cumulative distribution function giving the probability of an observed outcome having efficiencies relative to various objectives less than or equal to the efficiencies specified in the given M-tuple, i.e., $F(Z_t) = \text{Prob}(z_1 \leq z_{1t}, \ldots, z_M \leq z_{Mt})$.

II. Simplifications in the Statement of Loss Matrix Entries:

Because the statement above is too unwieldy for practical use, we generally make a series of simplifying assumptions. The first is that a_j and θ_n need not be considered explicitly since

everything of interest about them may be incorporated in the distribution function. With this simplification we may write:

$$l_{jn} = f(u(Z), F_{jn}(Z))$$

where $F_{jn}(Z)$ is the distribution function for given (a_j, θ_n).

Depressing a_j implies that the decision maker is not attracted to any action on grounds other than its performance relative to the stated objectives. If the list of objectives is incomplete such an assumption may fail. For example had we not considered the monetary cost of treatment in our objectives for the leprosy problem we could conceivably decide on a strategy which is optimal for the first three objectives but which is prohibitively expensive.

Incorporating θ_n in the distribution function assumes that the value of an outcome that can occur under any member of a set of states of nature is not altered by a specification of the particular member under which it did occur. In many problems the total set $\{\theta_n\}$ is such a set. In many others the set is more restricted because some objectives are simply not relevant under certain states of nature; this information can be incorporated in the distribution function. In the leprosy problem the minimization of new infections is clearly not a relevant objective when dealing with a non-contagious patient. This can be signified in the distribution function by assigning zero probability to all outcomes which specify that a new

infection appears under the state of nature corresponding to a non-contagious patient.

The second simplification which is generally made involves the assumption that the value and joint cumulative distribution functions are independent (i.e. the value of an outcome is not influenced by its probability of occurrence and vice versa), and the assumption that a mathematical expectation is the appropriate measure of effectiveness. Then we may write:

$$l_{jn} = E(u(Z)\ a_j, \theta_n) = \int_Z u(Z)dF_{jn}(Z)$$

where the expectation operation is to be interpreted as integration for continuous variables and summation for discrete ones. The use of a mathematical expectation will fail if the value function assigns an infinite value or cost to any possible outcome, i.e. one that occurs with non-zero probability.

In order to obtain a tractable statement of l_{jn} we usually attempt to make additional simplifying assumptions if they seem plausible for the problem at hand. With regard to the value function we often assume that the value of an outcome is the sum of the values of its components, i.e.

$$u(Z) = \sum_{m=1}^{m=M} u_m(z_m)$$

This assumption holds only if there is no interaction among the objectives in the sense that the value of a specific measure of

efficiency relative to one objective is not influenced by the measures of efficiency relative to other objectives which may be realized simultaneously. For example we could probably accept the supposition that the cost of having a lepromatous patient infect another individual is the same regardless of the duration of that patient's absence from his job. With this assumption we can write l_{jn} as

$$l_{jn} = \int_Z \sum_{m=1}^{m=M} u_m(z_m)dF_{jn}(Z) = \sum_{m=1}^{m=M} \int_Z u_m(z_m)dF_{jn}(Z) = \sum_{m=1}^{m=M} E(u_m(z_m) | a_j\theta_n)$$

In the same vein another simplification is often made based on the assumption that $u_m(z_m)$ is proportional to z_m, i.e. $u_m(z_m) = u_{m1}z_m$ where u_{m1} is a constant "unit value." With this assumption we can write the statement for the loss table entry as

$$l_{jn} = \sum_{m=1}^{m=M} u_{m1}E(z_m | a_j,\theta_n)$$

With regard to the joint cumulative distribution function, we can sometimes assume that the measures of efficiency are independent random variables. In this case we define an "efficiency function," $f_{jnm}(z_m)$ for each (a_j,θ_n,O_m) triplet.

1. For continuous quantitative objectives $(m = 1,\ldots,M_1)$, z_m is continuous and $f_{jnm}(z_m)$ is a probability density function.

2. For discrete quantitative objectives ($m = M_1+1, \ldots, M_2$), z_m is discrete and $f_{jnm}(z_m)$ is a probability distribution.
3. For qualitative objectives ($m = M_2+1, \ldots, M$) we usually take z_m to be 0 or 1 and $f_{jnm}(1)=f_{jnm}$ as the probability of achieving the objective O_m.

With this notation we may write

$$dF_{jn}(Z) = \prod_{m=1}^{m=M_1} f_{jnm}(z_m)dz_m \prod_{m=M_1+1}^{m=M} f_{jnm}(z_m)$$

Coupling this assumption to the preceding one simplifies computations because the mathematical expectations above are taken over univariate marginal distributions rather than over the multivariate joint distribution.

III. Some Methods for Obtaining Loss Matrix Entries: Few if any arts are poorer in rules than the art of selecting the proper value estimation technique for a particular problem, but at least we can say that the assumptions about the form of the value function inherent in a technique must be acceptable in the environment in which the technique is to be used. In this section we first consider methods which are applicable if the value function is a sum, and then consider a method which does not require that the value function be a sum.

Before considering specific methods we note that in applications we will attempt to impute value measures either from a

decision maker's past activities or from his responses in hypothetical decision situations. The results in the first case may fail to truly represent the decision maker's value structure in the environment where the solution is to be used because his past decisions may have been made in an environment where he had available less information about the implications of his acts than he will have at the conclusion of the study. Similarly his answers in hypothetical decision situations may be distorted by the lack of an immediate and real responsibility to act in such simulated environments. In either case the fact that we may be dubious of the derived values is one reason for an interest in sensitivity analysis since it will sometimes permit us to say how far "real" values can vary from estimated values before a strategy based on estimated values ceases to be optimal.

III A. Obtaining Loss Matrix Entries When the Value Function is a Sum: Here the procedures of Churchman et al. (1957) are applicable. In Churchman et al. (1957) and this section the assumption is made that $u(0,\ldots,0)=0$.

III A 1. Determining value functions when only qualitative objectives are present: In addition to the assumptions above, the procedure assumes that $\mathrm{Min}(z_m)=0$, $\mathrm{Max}(z_m)=1$, and $u_m = u_m(1) \geq 0$. In this case u_m is the only point needed for the value function of z_m. We will outline a procedure suitable for relatively few objectives (say M less than 8), referring to Churchman

et al. (1957) for an adaptation which is more suitable for a larger number.

a. The decision maker ranks the objectives in order of importance. We let O_1 denote the most and O_M denote the least important objectives.

b. The decision maker tentatively assigns 1.00 to u_1 and numbers in the range 0 to 1 to $u_2, \ldots, u_M$ reflecting his rough estimate of their importance relative to that of O_1.

c. We compare O_1 with $O_2+\ldots+O_M$ (where "+" denotes the logical connective "and").

(ci) If the decision maker responds that O_1 is less important than $O_2+\ldots+O_M$, we record $u_1 < u_2+\ldots+u_M$ and adjust (if necessary) the tentative estimates so that they satisfy this inequality. Then we repeat the procedure for O_1 versus $O_2+\ldots+O_{M-1}$, $O_2+\ldots+O_{M-2}$, etc. until a response of the form (cii) or (ciii) is obtained.

(cii) If the decision maker responds that O_1 and $O_2+\ldots+O_M$ are equally important, we record $u_1 = u_2 +\ldots+u_M$ and adjust (if necessary) the tentative estimates so that they satisfy this equation. Then we go to step d.

(ciii) If the decision maker responds that O_1 is more important than $O_2+\ldots+O_M$, we record $u_1 > u_2+\ldots+u_M$ and

adjust (if necessary) the tentative estimates so that they satisfy this inequality. Then we go to step d.

d. We compare O_2 with $O_3+\ldots+O_M$ proceeding as in step c, then compare O_3 with $O_4+\ldots+O_M$ etc. until the comparison of O_{M-2} with $O_{M-1}+O_M$ is completed.

e. We may then normalize the resulting estimates (i.e. require $\sum_{m=1}^{M} u_m = 1$) in order to express relative values, or in some cases we may choose to set $u_1 = 1$ (or 100) and express the other estimates as fractions (or percentages) of u_1.

An application of this technique to the specification of the relative importance of six types of defects which could appear in lots of a pharmaceutical product is reported by Stillson (1954). A board composed of nine men was presented with a series of questions of the sort, "Is a type 'a' defect more important than, less important than, or equally important as type 'b' and 'c' defects taken together?" In each case a vote was taken, the decision of the majority was accepted, and the tentative values assigned to the defects were adjusted to bring them into accord with all of the comparisons made. Letting u_1 through u_6 represent the importance of the six types of defects, we can condense the results given by Stillson (1954) into the twelve relations below.

1. $u_1 > u_2$ 4. $u_5 > u_6$ 7. $u_1 > u_2 + u_6$ 10. $u_3 < u_4 + u_6$
2. $u_2 > u_3$ 5. $u_6 > 0$ 8. $u_2 < u_3 + u_4 + u_5$ 11. $u_2 = u_3 + u_4$
3. $u_3 > u_4$ 6. $u_1 < u_2 + u_5$ 9. $u_2 > u_3 + u_5$ 12. $u_4 = u_5 + u_6$

We note that this list can be condensed further by discarding inequalities which are implied by other relations, e.g. 5 and 7 imply 1.

The decision was made to set $u_1 = 100$ and to express the other costs as percentages of u_1. A set of numbers satisfying these inequalities and equalities is (100, 80, 43, 37, 24, 13).

Having followed this procedure we have a set of numbers that can be used in the determination of an optimal strategy. A reasonable question to ask in such a situation is whether or not the same strategy would be optimal for some other set of numbers. In particular, since the given set resulted from a series of manipulations to make it fit into the system of linear equations and inequalities, we might ask if the same strategy would be optimal for any $(u_1, \ldots, u_M)$ satisfying the system.

III A 2. Determining the value functions when only quantitative objectives are present:

a. The first decision to be made is the selection of a value scale. Churchman et al. (1957) propose the following. When $z_1, \ldots, z_M$ are measured on the same scale use this as the value scale which involves the assumption that $u_m(z_m) = z_m$ for $m = 1, \ldots, M$. Otherwise transform $z_1, \ldots, z_M$ into a common scale which

(in order to lessen the effects of errors in the transformation process on the final solution) is taken to be the scale of efficiency, say z_1, of the "most important" objective which assumes that $u_1(z_1)=z_1$. We will follow this course in this section, but we note that problems may be encountered in which we will wish to have as our value scale one different than any scale available among our measures of efficiency. This may be a matter of convenience or may be because we have no z_m for which we are willing to assume $u_m(z_m)=z_m$.

As an example we may consider an inventory problem having as objectives the minimization of the average number of items carried in inventory measured in "items per week" and the minimization of the number of shortages (i.e. lost sales resulting from an exhaustion of stock) per week. We could construct a function that converts items in inventory per week to number of shortages per week or vice versa, or assign a dollar cost to the carrying of one item in inventory for a week and to a shortage so that the common scale is dollars per week.

Having selected z_1 as the measure of efficiency to serve as our value scale, we may prepare M-1 graphs having z_1 on the ordinate and $z_2,\ldots,z_M$ on the various abscisses. For each of these graphs we must determine correspondences between a sufficient number of points on the abscissa and points on the ordinate to plot a curve representing $u_m(z_m)$.

b. We may attempt to impute these correspondences by an examination of past decisions or by asking the decision maker questions of the type, "How many units relative to O_1 are as important to you as one unit relative to O_m?" Answers to such questions for various values of z_m and for $m = 2,\ldots,M$ will fix points on each graph to which a curve representing $u_m(z_m)$ may be fitted. A common procedure in curve fitting is to assume that the curve is representable by a polynomial

$$u_m(z_m) = \sum_{s=0}^{s=S_m} u_{ms} z_m^s$$

and to estimate the polynomial by least squares. In this connection we note that $u(0,\ldots,0)=0$ implies $u_{m0} = 0$ for all m and if $u_m(z_m)$ is proportional to z_m, $S_m = 1$.

As an example of the sort of questions that can arise following the use of this procedure, we suppose that we estimate that $u_2(z_2) = .6z_2$. Then we might ask, "How far can the unit value, u_{21}, vary on either side of .6 and still have the optimal strategy obtained for $u_{21} = .6$ remain optimal?" In addition we may want to know what other strategies become optimal as u_{21} shifts away from .6 in either direction.

c. In establishing the correspondences we may resort to a flanking attack, especially if the decision maker finds numerical responses difficult. For each graph we may select a pair of values (z_1^*, z_m^*) and ask, "Are z_1^* units relative to O_1 more

important than, less important than, or equally important as z_m^* units relative to O_m?" If the answer is "more," we repeat the question with z_1^{**} less than z_1^*, and if "less" we repeat it with z_1^{**} greater than z_1^*. We repeat the procedure until the range between the numbers z_1^* and z_1^{**} which bracket $u_m(z_m^*)$ is sufficiently small, and possibly duplicate the procedure for other points along the z_m axis. We could then fit a polynomial, $u_m^*(z_m)$, to the lower members of the pairs of points obtained and a polynomial, $u_m^{**}(z_m)$, to the upper members.

In some sense $u_m^*(z_m)$ and $u_m^{**}(z_m)$ "bracket" $u_m(z_m)$. We will take as the definition of "bracket" the statement that any $u_m(z_m)$ which is bracketed by $u_m^*(z_m)$ and $u_m^{**}(z_m)$ satisfies the relations:

$$\mathrm{Min}\left(\overset{(n)}{u_m^*}(z_m),\ \overset{(n)}{u_m^{**}}(z_m)\right) \leq \overset{(n)}{u_m}(z_m) \leq \mathrm{Max}\left(\overset{(n)}{u_m^*}(z_m),\ \overset{(n)}{u_m^{**}}(z_m)\right)$$

for $\mathrm{Min}\,(z_m) \leq z_m \leq \mathrm{Max}\,(z_m)$ and $n = 0, 1, 2, \ldots$
where $u_m^{(n)}(z_m)$ is the n-th derivative of the function and for a polynomial is given by:

$$u_m^{(n)}(z_m) = \sum_{s=n}^{s=G_m} (s)_n\, u_{ms}(z_m)^{s-n} \quad \text{where } (s)_n = s(s-1)\ldots(s-n+1)$$

Given an objective for which our information is of this type, a subject for analysis is whether or not there is a

unique optimal strategy for any $u_m(z_m)$ bracketed by the given functions. In part our ability to conduct such an analysis by the methods of subsequent chapters will depend on our ability to replace the relations defining "bracketing" by a system of linear constraints on the parameters of the value function. In the case of a polynomial function this can be done by examining the various derivatives from the highest order to the lowest, at each step replacing the general relation by an equivalent set of linear constraints on the coefficients of the polynomial.

(i) Thus we start with the S_m-th (or S-th for simplicity) derivative

$$u_m^{(S)}(z_m) = (S!)u_{mS}$$

and require

$$\mathrm{Min}\left(u_m^{\overset{(S)}{*}}(z_m),\ u_m^{\overset{(S)}{**}}(z_m)\right) \le u_m^{(S)}(z_m) \le \mathrm{Max}\left(u_m^{\overset{(S)}{*}}(z_m),\ u_m^{\overset{(S)}{**}}(z_m)\right)$$

for $\mathrm{Min}(z_m) \le z_m \le \mathrm{Max}(z_m)$, which is equivalent to:

$$\mathrm{Min}(u_{mS}^{*},\ u_{mS}^{**}) \le u_{mS} \le \mathrm{Max}(u_{mS}^{*},\ u_{mS}^{**})$$

For example we might obtain $u_{mS}^{*} \le u_{mS} \le u_{mS}^{**}$

(ii) We next examine the (S-1)-th derivative.

$$u_m^{(S-1)}(z_m) = ((S-1)!)\, u_{m\,S-1} + (S!)\, u_{mS}\, z_m$$

and require that:

$$\mathrm{Min}\left(u_m^{*(S-1)}(z_m),\ u_m^{**(S-1)}(z_m) \right) \leq u_m^{(S-1)}(z_m) \leq \mathrm{Max}\left(u_m^{*(S-1)}(z_m),\ u_m^{**(S-1)}(z_m) \right)$$

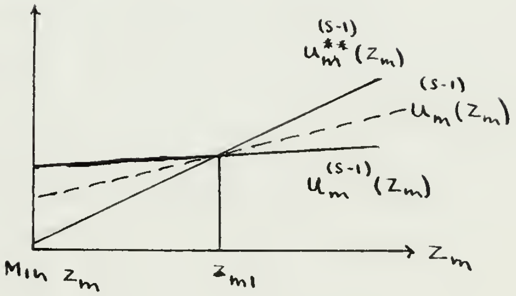
for $\mathrm{Min}(z_m) \leq z_m \leq \mathrm{Max}(z_m)$ which is equivalent to:

$$\mathrm{Min}(u^{*}_{m\,S-1} + S\, u^{*}_{mS} z_{mS},\ u^{**}_{m\,S-1} + S\, u^{**}_{mS} z_m) \leq u_{m\,S-1} + S\, u_{mS}\, z_m \leq \mathrm{Max}(\ldots)$$

for $\mathrm{Min}(z_m) \leq z_m \leq \mathrm{Max}(z_m)$.

The precise type of linear constraints which will replace this general statement of course depends on the nature of the problem. We will give two examples.

(1) Consider the graph on the right. In this case we can specify the desired relation by the two restrictions,



(a)

$$u^{**}_{m\,S-1} + S\, u^{**}_{mS}\, \mathrm{Min}(z_m) \leq u_{m\,S-1} + S\, u_{mS}\, \mathrm{Min}(z_m) \leq u^{*}_{m\,S-1} + S\, u_{mS}\, \mathrm{Min}(z_m)$$

and if $\mathrm{Min}(z_m) = 0$ this reduces to

$$u^{**}_{m\,S-1} \leq u_{m\,S-1} \leq u^{*}_{m\,S-1}$$

(b)

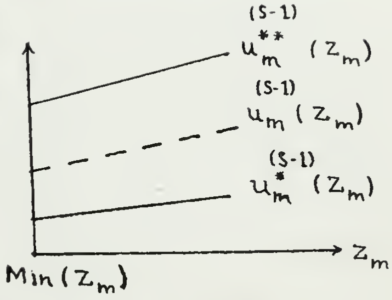
$$u_{m\,S-1} + S\,u_{mS}\,z_{m1} = \begin{array}{c} u^{*}_{m\,S-1} + S\,u^{*}_{mS}\,(z_{m1}) \\ \text{or} \\ u^{**}_{m\,S-1} + S\,u^{**}_{mS}\,(z_{m1}) \end{array}$$

We note that taken together these two constraints imply the constraint used to simplify (i).

(2) Now consider the second graph on the right. In this case we can specify the desired relation by:

(a) The constraint of (1)(a)

(b) The constraint of (i)



(iii) Next we look at the (S-2)-th derivative and find a set of linear constraints that together with the constraints already developed will insure that,

$$\mathrm{Min}\left(u_m^{*(S-2)}(z_m),\ u_m^{**(S-2)}(z_m) \right) \leq u_m^{(S-2)}(z_m) \leq \mathrm{Max}\left(u_m^{*(S-2)}(z_m),\ u_m^{**(S-2)}(z_m) \right)$$

for $\mathrm{Min}(z_m) \leq z_m \leq \mathrm{Max}(z_m)$

(iv) The process is repeated successively for all derivatives and the function itself, building a chain of linear constraints which will specify the type of variation which is permissible in the coefficients and the relations which must hold among them.

For a more concrete illustration of the procedure we suppose that in the inventory example noted earlier we decide to take the measure of efficiency relative to the objective of minimizing average inventory (O_1) as the value scale. With regard to O_2 (minimizing number of shortages) we decide that $\text{Min}(z_2) = 0$ and $\text{Max}(z_2) = 5$. We ask a series of questions like, "Is the carrying of 4 items in inventory for a month more costly than, less costly than, or as costly as incurring one shortage in a month?", and produce several pairs of points which we use to fit two cubics.

$$\left.\begin{aligned} u_2^{*}(z_2) &= z_2 + (z_2)^2 - .1(z_2)^3 \\ u_2^{**}(z_2) &= z_2 + 2(z_2)^2 - .2(z_2)^3 \end{aligned}\right. \qquad \text{for } 0 \leq z_2 \leq 5$$

These two equations bracket the value function

$$u_2(z_2) = u_{20} + u_{21}(z_2) + u_{22}(z_2)^2 + u_{23}(z_2)^3$$

To construct a system of linear constraints which will specify this bracketing relation, we first look at the third derivative and find that

$$\text{Min}\left(u_2^{*(3)}(z_2),\ u_2^{**(3)}(z_2) \right) \leq u_2^{(3)}(z_2) \leq \text{Max}\left(u_2^{*(3)}(z_2),\ u_2^{**(3)}(z_2) \right)$$

for $0 \leq z_2 \leq 5$ is equivalent to:

$$(1) \quad -.2 \leq u_{23} \leq -.1$$

Examining the second derivative we find that two linear constraints are equivalent to:

$$\text{Min}\ (\ u_2^{*}\ (z_2),\ u_2^{**}(z_2)\ ) \leq u_2\ (z_2) \leq \text{Max}\ (\ u_2^{*}\ (z_2),\ u_2^{**}(z_2)\ )\ \text{for}\ 0 \leq z_2 \leq 5$$

They are: $(2) \quad 1 \leq u_{22} \leq 2 \qquad (\text{setting}\ z_2 = 0)$

$(3) \quad u_{22} + 10\ u_{23} = 0 \qquad (\text{setting}\ z_2 = 10/3)$

Examining the first derivative we find that,

$$\text{Min}\ (\ u_2^{*}\ (z_2),\ u_2^{**}(z_2)\ ) \leq u_2\ (z_2) \leq \text{Max}\ (\ u_2^{*}(z_2),\ u_2^{**}(z_2)\ )\ \text{for}\ 0 \leq z_2 \leq 5$$

is equivalent to (1), (2), (3) and the following equation:

$$(4) \qquad u_{21} = 1 \qquad\qquad (\text{setting}\ z_2 = 0)$$

Examining the function itself we find that,

$$\text{Min}\ (\ u_2^{*}\ (z_2),\ u_2^{**}(z_2)\ ) \leq u_2(z_2) \leq \text{Max}\ (u_2^{*}(z_2),\ u_2^{**}(z_2))\ \ \text{for}\ 0 \leq z_2 \leq 5$$

is equivalent to (1) through (4) and,

$$(5) \qquad u_{20} = 0 \qquad\qquad (\text{setting}\ z_2 = 0)$$

We note that (2) and (3) imply (1). Therefore we can express the

fact that $u_2^*(z_2)$ and $u_2^{**}(z_2)$ bracket $u_2(z_2)$ by the system of linear constraints:

$$\begin{array}{lllrcl}
 & & u_{22} & -1 & \geq & 0 \\
 & & -u_{22} & 2 & \geq & 0 \\
 & & u_{22} + 10\, u_{23} & & = & 0 \\
 & u_{21} & & -1 & = & 0 \\
u_{20} & & & & = & 0
\end{array}$$

d. Another alternative for establishing the correspondences is suggested by Churchman et al. (1957) and involves application of the technique discussed in section III A 1 to a set of roughly comparable measures of efficiency $(z_1, \ldots, z_M)$.

For the t-th appraisement we may select some value of z_1, say z_{1t}, and for $m = 2, \ldots, M$ values z_{mt} such that we suspect that $u_m(z_{mt}) \leq z_{1t}$. We apply the technique to this set of numbers to determine weights $(w_{1t}, \ldots, w_{Mt})$ which can be used to establish points on all graphs. If in step c. of the procedure we set $w_{1t} = 1$ expressing $0 \leq w_{mt} \leq 1$ for $m = 2, \ldots, M$ as fractions of w_{1t}, the point on the ordinate of the (M-1)st graph corresponding to z_{mt} is $w_{mt} z_{mt}$.

T such appraisements will establish T points on each graph which we may use to plot a curve representing $u_m(z_m)$. These curves in turn can be used in determining an optimal strategy.

In such cases a question similar to the one that arose in section III A 1 may be of interest. Since the w_{mt}'s produced by the weighting procedures result from manipulations to make them satisfy systems of linear constraints, we may reasonably ask if there is a unique optimal strategy for all w_{mt}'s which satisfy the systems of linear constraints.

Certainly the (M-1) value functions will vary as the w_{mt}'s vary within the convex set defined by the systems of linear constraints, and in the case of general functions the manner in which the w_{mt}'s specify the parameters of the value functions is not obvious. However we note for future reference that if we employ the convention $w_{1t}=1$ and use polynomials fitted by least squares as our value functions, the coefficients of these polynomials will be linear functions of the w_{mt}'s. We will illustrate this for a quadratic.

We have made T appraisements and have determined the weights $(1,w_{21},\ldots,w_{M1}),\ldots,(1,w_{2T},\ldots,w_{MT})$. On the graph of z_m versus z_1 we have T points $(z_{mt}, w_{mt}z_{mt})$ for $t = 1,\ldots,T$. We refer to Graybill (1961) for a general discussion of least squares fitting of polynomials, and merely note here that if we assume,

$$u_m(z_m) = u_{m1}\, z_m + u_{m2}\, z_m^2$$

the coefficients of the quadratic fitted by least squares are the solutions of the system of equations:

$$\begin{pmatrix} \sum_{t=1}^{t=T} z_{mt}^2 & \sum_{t=1}^{t=T} z_{mt}^3 \\ \sum_{t=1}^{t=T} z_{mt}^3 & \sum_{t=1}^{t=T} z_{mt}^4 \end{pmatrix} \begin{pmatrix} u_{m1} \\ \\ u_{m2} \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{t=T} w_{mt} z_{mt}^2 \\ \sum_{t=1}^{t=T} w_{mt} z_{mt}^3 \end{pmatrix}$$

The solution of the system is:

$$\begin{pmatrix} u_{m1} \\ u_{m2} \end{pmatrix} = \frac{1}{D} \sum_{t=1}^{t=T} w_{mt} \begin{pmatrix} \sum_{t=1}^{t=T} z_{mt}^4 & - \sum_{t=1}^{t=T} z_{mt}^3 & z_{mt}^2 \\ - \sum_{t=1}^{t=T} z_{mt}^3 & \sum_{t=1}^{t=T} z_{mt}^2 & z_{mt}^3 \end{pmatrix}$$

where $D = (\sum_{t=1}^{t=T} z_{mt}^2)(\sum_{t=1}^{t=T} z_{mt}^4) - (\sum_{t=1}^{t=T} z_{mt}^3)^2$ is the determinant of the first 2x2 matrix above.

We note that the requirement $w_{mt} = 1$ for $m = 2,\dots,M$ is not absolute. We try to pick values z_{mt} such that $u_m(z_{mt}) = z_{1t}$ in order to insure that the convex set defined by the linear constraints on $(w_{2t},\dots,w_{Mt})$ is closed. This point will be discussed in detail in Chapter 3.

III A 3. Determining value functions when both qualitative and quantitative objectives are present: In this case we may use the procedures of section III A 2 to determine $u_m(z_m)$ for the quantitative objectives $(m = 1,\dots,M_2)$ and the procedure suggested by Churchman et al. (1957) to determine u_m for the

qualitative objectives ($m = M_2 + 1, \ldots, M$). This procedure consists of:

a. Select a subset of quantitative objectives, say $O_1, \ldots, O_t, \ldots, O_T$ ($T \leq M_2$).

b. For each O_m, O_t pair ($m = M_2+1, \ldots, M$; $t = 1, \ldots, T$) ask a question like, "How many units of achievement relative to O_t are as important as the attainment of O_m?" The answer specifies a number, z_{tm}, on the z_t scale.

c. For each O_m ($m = M_2+1, \ldots, M$) use the value functions $u_t(z_t)$ for $t = 1, \ldots, T$ to transform the T results of step b into the standard value scale.

We may take one of the results of step c or some combination of them as representing u_m. In addition we may be interested in viewing the minimum and maximum of the results of step c as representing lower and upper bounds on u_m and in questioning the uniqueness of an optimal strategy for any set $(u_{M_2+1}, \ldots, u_M)$ satisfying these restrictions. Of course if the range between the maximum and minimum for some u_m is excessive, we may question the consistency of the decision maker's responses.

An alternative to step b is to attempt to bracket the z_t which is equivalent to the attainment of O_m with questions like, "Is the attainment of z_t units relative to O_t more important than, less important than, or equally important as the attainment of O_m?" The bracketing values, $z^*_{tm} \leq z_{tm} \leq z^{**}_{tm}$, can be transformed

in step c, and we may take as lower and upper bounds on u_m:

$$\underset{t}{\mathrm{Max}}\; u_t\,(z^{*}_{tm}) \leq u_m \leq \underset{t}{\mathrm{Min}}\; u_t\,(z^{**}_{tm})$$

Then we select a set $(u_{K_2+1},\ldots,u_M)$ satisfying the constraints and use it to determine an optimal strategy, or we may attempt to determine if there is a unique optimal strategy for any set which satisfies the constraints.

III A 4. Determining the loss matrix entries for a given value function: In this section we present only one of the several methods which may be used and refer to Churchman et al. (1957) for a discussion of the others.

a. Determining the relevant probability distributions: We will not attempt to cover even the simplest aspects of the estimation of probability distributions, but will assume that by review of historical data or by suitable experiments, we can record the outcomes which resulted in many applications of a_j under θ_n for $j = 1,\ldots,J$ and $n = 1,\ldots,N$. We note that in practice our inability to observe an adequate number of outcomes for each (a_j,θ_n) pair may be a major hindrance in the way of determining the loss matrix precisely.

(1) Under the assumption that the joint probability distribution is not factorable, we compute the relative frequency of occurrence of each M-tuple (Z_t) when a_j is applied under θ_n, and take this as out estimate of $dF_{jn}(Z_t)$ for each M-tuple.

(ii) Under the assumption that the joint probability distribution is factorable, we compute the relative frequency of occurrence of each value of $z_m (m=1,\ldots,M)$ when a_j is applied under θ_n, and take this as our estimate of $f_{jnm}(z_m)$ for each value of z_m.

The use of empirical histograms requires that we divide the z_m scale for continuous quantitative objectives into class intervals, each interval being regarded as a value assumed by the measure of efficiency for a discrete quantitative objective.

b. Determining the measure of effectiveness: We first transform all of the measures of efficiency observed in the data of a. to the standard scale by using the value functions $u_m(z_m)$.

(i) If the joint probability distribution is not factorable, we compute the value of each observed outcome as:

$$u(Z_t) = \sum_{m=1}^{m=M} u_m(z_{mt})$$

and then compute the loss matrix entry (measure of effectiveness) by

$$l_{jn} = E\,(u(Z) \mid a_j, \theta_n) = \sum_{Z} u(Z_t)\, dF_{jn}(Z_t)$$

where $dF_{jn}(Z_t)$ is from a(i) above.

(ii) If the joint probability distribution is factorable, we compute for each objective

$$E\,(u_m(z_m) \mid a_j, \theta_n) = \sum_{z_m} u_m(z_m)\, f_{jnm}(z_m)$$

where $f_{jnm}(z_m)$ is from a(ii) above, and then sum these numbers to obtain

$$l_{jn} = \sum_{m=1}^{m=M} E\left(u_m(z_m) \mid a_j, \theta_n\right)$$

III B. Obtaining Loss Matrix Entries When the Value Function is not a Sum:

III B 1. Determining the value function: When the value of an outcome cannot be expressed as the sum of value functions on the M measures of efficiency which define the outcome, we must resort to methods for estimating a value function over a set of mutually exclusive events. The mutually exclusive events to which we apply such methods will be either the outcomes or the events "a_j is adopted in the presence of θ_n." In the latter case we would be estimating the loss matrix entries directly, informally evaluating the implications of adopting a_j under θ_n. These methods are equally applicable when the value function is a sum, but the methods discussed in section III A usually are to be preferred in that case.

In this section we will discuss only the most popular method for estimating a value function over a set of mutually exclusive events. It is based on the von Neumann-Morgenstern theory of utility, and its philosophical basis and practicability have been the subjects of many inquiries (see e.g. Churchman (1961) or Luce and Raiffa (1957)). The names attached

to and variations on the method are legion; we will call it the "standard gamble" and will present only one version which can be described as follows letting Z_t denote the outcome $(z_{1t},\ldots,z_{mt})$ and $u(Z_t)$ denote $u(z_{1t},\ldots,z_{mt})$.

We rank the outcomes and let $1 = u(Z_T) = \text{Max } u(Z_t)$ and $0 = u(Z_1) = \text{Min } u(Z_t)$. The specification of the maximum and minimum entails no loss of generality because it is only necessary to have the loss matrix (and consequently the value function) unique up to a positive linear transformation (see examples 1 and 3 of chapter 4).

For each other outcome, Z_t, we consider a series of hypothetical gambling situations in which the decision maker must choose between the occurrence of Z_t for sure and a gamble for which outcome Z_T will occur with probability p_t and outcome Z_1 will occur with probability $(1-p_t)$. For each such situation the decision maker must choose between the value $u(Z_t)$ for sure and the expected value,

$$(1 - p_t)\, u(Z_1) + p_t\, u(Z_T) = (1 - p_t)(0) + p_t(1) = p_t$$

of a gamble. If he chooses Z_t we impute that $u(Z_t)$ is greater than p_t. If he chooses the gamble we impute that $u(Z_t)$ is less than p_t. If he is indifferent as to his choice we impute that $u(Z_t) = p_t$. We vary the probability, p_t, used from trial to trial in order to conclude this series with two numbers p_t^* and p_t^{**} such that, T

$$0 \le p_t^* \le u(Z_t) \le p_t^{**} \le 1$$

Of course if $p_t^* = p_t^{**}$ we will have determined $u(Z_t)$ precisely. We note that an important assumption of this method is that the decision maker is neither attracted or repelled by the prospect of gambling in itself.

If there are a finite number, T, of outcomes we will conclude this procedure with the results $u(Z_1)=0$, $p_t^* \le u(Z_t) \le p_t^{**}$ for t = 2,...., T-1, and $u(Z_T)=1$. To obtain a loss matrix we may take as $u(Z_t)$ some number such as $(1/2)(p_t^* + p_t^{**})$ satisfying the given constraints for t = 2,...,T-1. However, we may also want to determine if there is a unique optimal strategy for any set $u(Z_2),\ldots,u(Z_{T-1})$ satisfying the constraints.

If some z_m used to define the outcomes is continuous, the standard gamble cannot be used for all Z_t. Either we must divide the z_m scale into class intervals each of which is regarded as a discrete value in order to obtain results like those of the preceding paragraph or we must select a suitable finite subset of outcomes (which includes Z_1 and Z_T), apply the standard gamble procedure to these outcomes, and then fit a function or functions to these results in order to be able to find the value of any outcome.

The latter procedure may be desirable even when all measures of efficiency are discrete if the number of outcomes is so large as to make impractical the application of the standard

gamble procedure to every outcome. If the procedure is used we might again take a polynomial to be the value function. Such a polynomial would be of the form:

$$u\ (Z) = u_o + \sum_{m=1}^{m=M} u_m\ z_m + \sum_{m=1}^{m=M}\ \sum_{i=1}^{i=M}\ u_{mi}\ z_m\ z_i + \ldots$$

If the polynomial is a cubic or of higher order and M is greater than two or three, the fitting of such a function will involve a large number of computations. However presuming it is computationally feasible, we could fit one polynomial, $u^*(Z)$, using the p_t^* values obtained in the standard gamble situations and another, $u^{**}(Z)$, using the p_t^{**} values. These two multivariate polynomials may be viewed as bracketing $u(Z)$ where "bracketing" is defined as in section III A 2c for univariate functions but in terms of all partial derivatives. In a manner analogous to that described for univariate polynomials, we can reduce the bracketing definition to a system of linear constraints on the coefficients of the polynomial. We may use some polynomial satisfying these constraints to obtain the loss matrix or we may choose to ask if there is a unique optimal solution for any polynomial value function bracketed by $u^*(Z)$ and $u^{**}(Z)$.

III B 2. Determining the loss matrix entries for a given value function: The discussion of section III A 4a on the determination of the relevant probability distribution is applicable here. Having obtained the value function and

probability distribution, we compute the loss matrix entry (measure of effectiveness) as:

$$l_{jn} = \sum_{Z} u\,(Z_t)\; dF_{jn}\,(Z_t)$$

III B 3. Application of the standard gamble directly to the loss matrix: The application of the standard gamble directly to the (a_j, θ_n) pairs proceeds in essentially the manner described in section III B 1. For notational convenience we let $l_{JN} = 1$ represent $\mathrm{Max}(l_{jn})$ and $l_{11} = 0$ represent $\mathrm{Min}(l_{jn})$. We then use situations involving on the one hand l_{jn} for sure and on the other hand a gamble with the expectation

$$(1 - p_{jn})\, l_{11} + p_{jn}\, l_{JN} = p_{jn}$$

For each l_{jn} we find two numbers, p^{*}_{jn} and p^{**}_{jn}, such that

$$0 \le p^{*}_{jn} \le l_{jn} \le p^{**}_{jn} \le 1 \quad \text{for } j = 1,\ldots,J \text{ and } n = 1,\ldots,N$$

We may take as our loss matrix entries some numbers which satisfy the constraints above; for example we could let

$$l_{jn} = (1/2)(p^{*}_{jn} + p^{**}_{jn}) \quad \text{for } j = 1,\ldots,J \text{ and } n = 1,\ldots,N$$

However we may also wish to ask if there is a strategy for the statistical decision problem which is optimal for every loss matrix containing elements that satisfy the constraints generated by the standard gamble procedure.

IV. Summary

IV A. Assumptions of Sensitivity Analysis Methods: In chapters 3 and 4 we will present methods for analyzing the sensitivity of the optimal strategy in a finite statistical decision theory problem. The methods are designed to handle parameters that enter the criterion matrix linearly and are subject to linear constraints. That is the methods are suitable for parameters $(u_1,\ldots,u_M)$ such that

1. The criterion matrix can be expressed

$$C = C_o + \sum_{m=1}^{m=M} u_m C_m$$

where C_m is an arbitrary JxI matrix for $m = 0,1,\ldots,M$

2. All restrictions on and relations among the parameters can be expressed by a system of linear constraints

$$A_1^0 U + B_1^0 \overset{>}{(\geq)} 0$$

$$A_2^0 U + B_2^0 = 0$$

where $U' = (u_1,\ldots,u_M)$ is the vector of parameters, $A_1^0(RxM)$ and $A_2^0(PxM)$ are coefficient matrices, and $B_1^0(Rx1)$ and B_2^0 $(Px1)$ are vectors of constants. The notation $\overset{>}{(\geq)}$ means that both strict and ordinary inequalities may be present.

3. When M is greater than 1 we also assume that the convex set defined by the system of linear constraints is bounded,

i.e. the set can be enclosed by an hypersphere of finite radius. This assumption will be discussed in Chapter 3.

Condition 1 may arise in a variety of ways. In this section and the examples of Chapters 3 and 4 we will consider cases in which the condition obtains as a result of the fact that the loss matrix can be expressed

$$L = E_o + \sum_{m=1}^{m=M} u_m E_m$$

where $E_m (m = 0,1,\ldots,M)$ is an arbitrary JxN matrix. Then we may write

$$C = L\Phi P = E_o\Phi P + \sum_{m=1}^{m=M} u_m E_m \Phi P = C_o + \sum_{m=1}^{m=M} u_m C_m$$

That is by setting $C_m = E_m \Phi P (m = 0,1,\ldots,M)$ we deduce condition 1. from the condition that the loss matrix be expressable as a linear function of the M parameters under consideration. In Chapter 5 we will consider how condition 1. may arise in a context other than the study of parameters entering the loss matrix.

IV B. Review of Value Estimation Methods: To exemplify the manner in which the above conditions may arise, we will review the methods discussed in this chapter showing how the questions suggested earlier are equivalent to questions about the sensitivity of the optimal solution to parameters satisfying those conditions.

When the value function is a sum (section III A) we can express the elements of the loss matrix,

$$l_{jn} = \sum_{m=1}^{m=M} E\,(u_m\,(z_m) \mid a_j,\, \theta_n)$$

where the expectation may be over a joint or marginal probability distribution.

When only qualitative objectives are present (section III A 1),

$$E\,(u_m\,(z_m) \mid a_j,\, \theta_n) \;=\; u_m\,E\,(z_m \mid a_j,\, \theta_n) \quad \text{where } z_m = 0 \text{ or } 1$$

If we let E_m have elements

$$e_{mjn} = E\,(z_m \mid a_j,\, \theta_n) \quad \text{for } j = 1,\ldots,J\,;\; n = 1,\ldots,N$$

we can express the loss matrix as,

$$L = \sum_{m=1}^{m=M} u_m\,E_m$$

In addition the restrictions on and relations among $u_1,\ldots,u_M$ as determined by the Churchman-Ackoff procedure are summarized by linear equations and strict inequalities, and the combination of non-negativity assumptions with normalization (step e of the procedure) insures that the convex set is bounded. Therefore the methods of Chapters 3 and 4 may be used to determine if there is a unique optimal strategy for any such set $u_1,\ldots,u_M$, and if there is not to delineate the subsets $u_1,\ldots,u_M$ for which particular strategies are optimal.

When only quantitative objectives are present (section III A 2) and we represent the value functions by polynomials, we have,

$$E\,(u_m(z_m)\mid a_j,\,\theta_n) = E\,(\sum_{s=0}^{s=S_m} u_{ms}\, z_m^s \mid a_j,\,\theta_n) = \sum_{s=0}^{s=S_m} u_{ms}\, E(z_m^s \mid a_j,\,\theta_n)$$

and the elements of the loss matrix are given by,

$$l_{jn} = \sum_{m=1}^{m=M} \sum_{s=0}^{s=S_m} u_{ms}\, E\,(z_m^s \mid a_j,\,\theta_n)$$

If we let E_{ms} have elements $e_{msjn} = E(z_m^s \mid a_j,\,\theta_n)$ for $j = 1,\ldots,J$ and $n = 1,\ldots,N$ we can express the loss matrix as,

$$L = \sum_{m=1}^{m=M} \sum_{s=0}^{s=S_m} u_{ms}\, E_{ms}$$

In particular the contribution to this sum for an m such that $u_m(z_m)$ is proportional to z_m is $u_{m1}E_{m1}$. In section III A 2c we showed for polynomial value functions, that the "bracketing" relations can be reduced to a system of equality and ordinary inequality constraints on the coefficients u_{ms} above. Furthermore the most lax form taken by such a system of constraints is $u_{ms}^{*} \leq u_{ms} \leq u_{ms}^{**}$ for $s = 0,\ldots,S_m$ and $m = 1,\ldots,M$ which defines a bounded convex set. Therefore we are in a position to determine if there is a unique optimal strategy for a set of M polynomial value functions bracketed by a set of M pairs of polynomials, and if there is not we may specify the subsets

$\{u_{ms}\}$ $s = 1,\ldots,S_m; m = 1,\ldots,M$ for which particular strategies are optimal.

In section III A 2d we presented a method for determining the coefficients of the polynomial value functions as linear functions of MT weights, $w_{mt}(m = 1,\ldots,M; t = 1,\ldots,T)$. If we express these linear relations by

$$u_{ms} = \sum_{t=1}^{t=T} a_{mst} \; w_{mt}$$

where a_{mst} is a constant determined by the least squares procedure, we can state the loss matrix as,

$$L = \sum_{m=1}^{m=M} \sum_{s=1}^{s=S_m} \sum_{t=1}^{t=T} a_{mst} \; w_{mt} \; \ell_{ms} = \sum_{m=1}^{m=M} \sum_{t=1}^{t=T} w_{mt} \sum_{s=1}^{s=S_m} \ell_{mst}$$

$$\text{where } \ell_{mst} = a_{mst} \; \ell_{ms}$$

The w_{mt}'s are derived by the Churchman-Ackoff procedure so the restrictions on and relations among them are summarized by linear equations and strict inequalities. For each of the T appraisements in section III A 2d we tried to pick values z_{mt} such that the constraints on the w_{mt}'s would contain constraints equivalent to $w_{mt} \leq 1$ in which case the convex set will certainly be bounded. Otherwise, depending on the particular constraint, the set may be unbounded so that one or more w_{mt} is not restricted to finite values. Roughly speaking each w_{mt} fixes the point $u_m(z_{mt})$ and we do not expect to encounter a value function which can assume infinite values for

finite values of the measures of efficiency. Therefore it seems reasonable in this case to regard an unbounded set as a failure of our estimation procedure against which we should incorporate safeguards in our analysis procedure. This will be done in the method proposed in Chapter 3 for the determination of whether or not there is a unique optimal solution for any set $w_{11},\ldots,w_{1T},\ldots,w_{N1},\ldots,w_{NT}$ satisfying the constraints.

In section III B we considered the standard gamble method for obtaining estimates of the value of outcomes. If there are T possible outcomes $Z_1,\ldots,Z_T$, we may let $u_t(Z_t) = u_t(u_1 = 0, u_T = 1)$ and E_t be a matrix with entries $dF_{jn}(Z_t)$ for $j = 1,\ldots,J$ and $n = 1,\ldots,N$. We can express the loss matrix as:

$$L = \sum_{t=1}^{t=T} u_t E_t = E_T + \sum_{t=2}^{T-1} u_t E_t$$

Because the set $u_2,\ldots,u_{T-1}$ is subject only to linear restrictions, we are in a position to determine if there is a unique optimal strategy for any set such that $p_2^* \leq u_2 \leq p_2^{**},\ldots,p_{T-1}^* \leq u_{T-1} \leq p_{T-1}^{**}$ which is bounded and if there is not we may specify the subsets $u_2,\ldots,u_{T-1}$ for which particular strategies are optimal.

If we have chosen to use a multivariate polynomial as the value function on Z, we may let E_0 be a matrix of ones, E_m have elements $E(z_m \mid a_j,\theta_n)$. E_{mi} have elements $E(z_m z_i \mid a_j,\theta_n)$, etc.

Then we may express the loss matrix as,

$$L = u_o K_o + \sum_{m=1}^{m=M} u_m K_m + \sum_{m=1}^{m=M} \sum_{i=1}^{i=M} u_{mi} K_{mi} + \cdots$$

As was noted for the bracketing relation on a univariate polynomial, we may use the methods of Chapters 3 and 4 to determine if there is a unique optimal strategy for any polynomial value function bracketed by the polynomials $u^*(Z)$ and $u^{**}(Z)$, and if there is not we may specify the subsets $u_0, u_1, \ldots, u_M, u_{11}, \ldots, u_{MM}$, etc. for which particular strategies are optimal.

If we apply the standard gamble method directly to the (a_j, θ_n) pairs (section III B 3) and have obtained estimates of the loss matrix entries, we can let E_{jn} have a one as its (jn) element and zeros elsewhere. Then the loss matrix may be expressed as:

$$L = \sum_{j=1}^{j=J} \sum_{n=1}^{n=N} l_{jn} E_{jn}$$

Because the standard gamble procedure subjects each l_{jn} to only linear restrictions (i.e. $l_{11} = 0$, $l_{JN} = 1$, $p^*_{jn} \leq l_{jn} \leq p^{**}_{jn}$) and determines a bounded convex set, we are in a position to determine if there is a unique optimal strategy for any loss matrix having entries that satisfy these restrictions, and if there is not to specify subsets of loss matrices for which particular strategies are optimal.

To illustrate the possibility of the failure of the conditions for sensitivity analysis, we consider a situation which could conceivably arise in the application of the methods of section III A 3 when the value function is a sum and both quantitative and qualitative objectives are present. Say in step a we select as the quantitative objective O_T one for which we have determined the value function to be $u_T(z_T) = u_{T1} z_T$. In the alternative form of step b for the qualitative objective O_M we find that $z^{*}_{TM} \leq z_{TM} \leq z^{**}_{TM}$, and in step c we transform these numbers into the common scale using $u_T(z_T)$ and find that,

$$k_1 = u_{T1}\, z^{*}_{TM} = \operatorname*{Max}_{t} \left(u_t(z^{*}_{tM}) \right) \geq k_3 = \operatorname*{Max}_{t \leq T-1} \left(u_t(z^{*}_{tM}) \right)$$

$$k_2 = u_{T1}\, z^{**}_{TM} = \operatorname*{Min}_{t} \left(u_t(z^{**}_{tM}) \right) \leq k_4 = \operatorname*{Min}_{t \leq T-1} \left(u_t(z^{**}_{tM}) \right)$$

Therefore we conclude that $k_1 \leq u_M \leq k_2$, and form the loss matrix,

$$L = E + u_{T1}\, E_{T1} + u_M\, E_M$$

where E_{T1} has as elements $E(z_T \mid a_j, \theta_n)$
E_M has as elements $E(z_M \mid a_j, \theta_n)$
E summarizes the contribution of all other objectives to the loss matrix.

Both u_{T1} and u_M are suitable parameters for sensitivity analysis by condition (i), but if we were to analyze the sensitivity of the optimal solution to variations in u_{T1} and u_M subject to the

constraints that $k_5 \leq u_{T1} \leq k_6$ and $k_1 \leq u_M \leq k_2$ we would be violating condition (2). Although the given constraints are linear, they do not express all of the restrictions on and relations between u_{T1} and u_M. These are given by

$$\mathrm{Max}\ (k_3,\ (z_{TM}^{*})\ u_{T1}) \leq u_M \leq \mathrm{Min}\ (k_4,\ (z_{TM}^{**})\ u_{T1})$$

$$\mathrm{Max}\ (k_5,\ \frac{k_3}{z_{TM}^{**}}) \leq u_{T1} \leq \mathrm{Min}\ (k_6,\ \frac{k_4}{z_{TM}^{*}})$$

The last inequalities presume positive z_{TM}^{*} and z_{TM}^{**}, and are included to prevent an inconsistency in the specification of u_M. The most straightforward way to avoid such a situation is to insure that no $u_m(z_m)$ which will be the subject of sensitivity analysis is included among the T objectives selected in step a of the procedure.

Having cautioned against the misuse of the methods of sensitivity analysis, we reiterate that these methods are not dependent on the use of the particular value estimation procedures presented in this chapter. In fact no value estimation procedure of any kind need be involved because the sensitivity analysis methods are valid for parameters which satisfy the given conditions regardless of the meaning attached to those parameters in a particular problem.

Chapter 3

Sensitivity Analysis for Variations in Several Parameters

In this chapter we turn to the formal procedure for analyzing the sensitivity of the optimal strategy of a finite statistical decision theory problem to variations in several parameters satisfying the conditions given in section IV A of Chapter 2.

We first observe that if a system of equality constraints is consistent, each nonredundant equation can be used to express one of the parameters as a linear function of the others. The procedure for accomplishing this is discussed in section I of Appendix I. Because it can be done, we can simplify the statement of the loss matrix. For example if we can write u_M as

$$u_M = d_o + \sum_{m=1}^{M-1} d_m u_m$$

we can express the criterion matrix as

$$C = C_o + \sum_{m=1}^{M-1} u_m C_m + u_M C_M = C_o + \sum_{m=1}^{M-1} u_m C_m + (d_o + \sum_{m=1}^{M-1} d_m u_m) C_M$$

$$C = C_o + d_o C_M + \sum_{m=1}^{M-1} u_m (C_m + d_m C_M) = C_o^* + \sum_{m=1}^{M-1} u_m C_m^*$$

and the equation from which we obtained u_M as a linear function of $u_1, \ldots, u_{M-1}$ no longer need be included among the constraints.

For ease of discussion we assume that the procedure of section I of Appendix I has been used to reduce the number of parameters considered explicitly and to eliminate the equality constraints. With this simplification we write the conditions for the applicability of the methods of analysis presented in this chapter as

1. The criterion matrix can be expressed

$$C = C_o + \sum_{m=1}^{m=M} u_m C_m$$

where C_m is an arbitrary JxI matrix for m = 1,...,M

2. All restrictions on and relations among the parameters can be expressed by a system of linear inequalities

$$AU + B \;(\geq)\; 0$$

where $U' = (u_1,\ldots,u_M)$ is the vector of parameters, A(RxM) is the coefficient matrix, and B(Rx1) is the vector of constants. The convex set defined by this system is assumed to be bounded.

When the criterion matrix is expressable as in 1. above, the expected loss for a strategy a_k is a linear function of U.

$$R\,(a_k) = \mathrm{tr}\,(S_kC) = \mathrm{tr}\,(S_kC_o) + \sum_{m=1}^{m=M} u_m\,\mathrm{tr}\,(S_kC_m)$$

I. *Investigation of Strategies Optimal for Given Constraints on Parameters:* In this section we consider a method for answering questions such as those asked in Chapter 2, i.e. given a system of linear constraints defining a convex set in which U must lie we will investigate the various strategies that are optimal for some admissible values of U.

I A. General Discussion: In general we expect that if U varies sufficiently over M-space, the optimal strategy derived from C will change. However the U with which we are concerned cannot range arbitrarily over M-space; it is restricted to variation within the bounded convex set defined by the system of linear inequalities in 2. above. In order to determine if there is a strategy which is optimal for any U in such a convex set, we may employ the following result,

> If a strategy is optimal at all of the extreme points of a bounded convex set, it is optimal at every point in the set.

where an extreme point is defined as follows.

> A point U is an extreme point of a convex set if and only if there do not exist other points U^* and U^{**} $(U^* \neq U^{**})$ in the set such that

$$U = v\,U^* + (1-v)\,U^{**} \qquad 0 < v < 1$$

To prove this statement we first note that if a strategy is optimal at two points in M-space it is optimal for any

convex combination of those two points. Consider the points U^* and U^{**} and assume that there exists a strategy s_o which is optimal at both points. That is for any strategy s_k

$$\mathrm{tr}\,(S_oC^*) \leq \mathrm{tr}\,(S_kC^*) \text{ and } \mathrm{tr}\,(S_oC^{**}) \leq \mathrm{tr}\,(S_kC^{**})$$

where

$$C^* = C_o + \sum_{m=1}^{m=M} u_m^* C_m \quad \text{and} \quad C^{**} = C_o + \sum_{m=1}^{m=M} u_m^{**} C_m$$

Then for any v such that $0 \leq v \leq 1$ we have

$$v\,\mathrm{tr}\,(S_oC^*) \leq v\,\mathrm{tr}\,(S_kC^*) \text{ and } (1-v)\,\mathrm{tr}\,(S_oC^{**}) \leq (1-v)\,\mathrm{tr}\,(S_kC^{**})$$

The sum of these two results yields

$$v\,\mathrm{tr}\,(S_oC^*) + (1-v)\,\mathrm{tr}\,(S_oC^{**}) \leq v\,\mathrm{tr}\,(S_kC^*) + (1-v)\,\mathrm{tr}\,(S_kC^{**})$$

$$\mathrm{tr}\,(S_o\,[\,v\,C^* + (1-v)\,C^{**}\,]\,) \leq \mathrm{tr}\,(S_k\,[\,v\,C^* + (1-v)\,C^{**}])$$

Therefore s_o is an optimal strategy for any criterion matrix which can be expressed as

$$C = v\,C^* + (1-v)\,C^{**} \qquad \text{where } 0 \leq v \leq 1$$

Now consider a point U which is a convex combination

$$U = v\,U^* + (1-v)\,U^{**} \qquad \text{where } 0 \leq v \leq 1$$

The criterion matrix corresponding to U is

$$C = C_o + \sum_{m=1}^{m=M} u_mC_m = v\,C_o + (1-v)\,C_o + \sum_{m=1}^{m=M} (v\,u_m^* + (1-v)\,u_m^{**})\,C_m$$

$$= v\,(C_o + \sum_{m=1}^{M-1} u_m^{*} C_m) + (1-v)\,(C_o + \sum_{m=1}^{M-1} u_m^{**} C_m) = v\,C^{*} + (1-v)\,C^{**}$$

and this is precisely the form of criterion matrix for which s_o is optimal. Thus s_o is an optimal strategy at every convex combination of U^{*} and U^{**}.

An analogous proof would provide the same result if we were to consider the convex combinations of an arbitrary finite number of points. Every point of a bounded convex set defined by linear constraints can be expressed as a convex combination of a finite number of "extreme points," the set of extreme points being a subset of the set of $\binom{R}{M}$ points obtained by letting M constraints at a time hold as equalities. Therefore if a strategy is optimal at every extreme point of a bounded convex set, it is optimal at every point of the set.

We have sacrificed some generality for the convenience of expressing every point in the convex set as a convex combination of the extreme points of the set. The assumption of a bounded convex set implies that we can at least specify a very large positive number and a very large negative number between which any parameter under scrutiny can be considered to lie. As a rule these numbers need not be specified because their inclusion in the analysis is obviated by the other constraints. In any case we need do no more in our formulation than specify such numbers for parameters that are obviously not otherwise constrained to finite values. The analysis procedure will

signal any unboundedness of the convex set, and so we may defer the specification of these numbers for other parameters until they are shown to be necessary.

Because the minimal (in the sense of requiring the fewest constraints) bounded convex set defined by linear inequalities in M-space is the "simplex" defined by M+1 constraints, a necessary condition for the convex set to be bounded is $R \geqq M+1$. Therefore we will assume that we have more linear inequality constraints than parameters.

In order to determine if there is a strategy which is optimal for all U satisfying $AU + B \overset{>}{(\geqq)} 0$, we may generate the extreme points of the convex set defined by the linear inequalities, compute a criterion matrix for each extreme point, and determine the corresponding optimal strategies. If one strategy is optimal at every extreme point, it is optimal over the set.

In Appendix I we present a procedure for determining the extreme points of a convex set defined by linear inequalities. It provides a method for determining a first extreme point solution and positive indications of inconsistency of the system or unboundedness of the convex set. The procedure provides a method for moving from the first extreme point to "adjacent" extreme points which in turn are used to find other extreme points.

We compute C_m for $m = 0,1,\ldots,M$ and having obtained the extreme points of the convex set, we compute the criterion matrix for each extreme point by

$$C = C_o + \sum_{m=1}^{M} u_m C_m$$

The corresponding optimal strategies are obtained by the procedure outlined in Chapter 1. It is possible that more than one strategy is optimal at an extreme point in which case all optimal strategies should be noted. After this procedure has been followed for all extreme points if one strategy has been among the optimal strategies at each extreme point it is optimal over the set.

If we find that no one strategy is optimal at every extreme point, we may wish to tabulate all strategies which are candidates for the optimal strategy if we know no more about U than the fact that it satisfies all inequalities and to specify the subsets within which each such strategy is optimal. Say we have V strategies $s_1,\ldots,s_V$ each of which is optimal at one or more extreme points. For any two such strategies, say s_v and s_w, we can specify the hyperplane in M-space on which the expected loss for the two strategies is equal by

$$0 = E\,(s_v) - E\,(s_w) = tr\,((S_v-S_w)C) = tr\,((S_v-S_w)C_o)+\sum_{m=1}^{M} u_m\, tr((S_v-S_w)C_m)$$

$$0 = A_{vw}\,(U)$$

or by

$$0 = E\,(s_w) - E\,(s_v) = tr\,((S_w-S_v)\,C)$$

$$0 = A_{vw}(U) = -A_{vw}(U)$$

for v, w = 1,..., V and $v \neq w$.

The two forms specify the same hyperplane in M-space and are equivalent for our purposes. There are $V(V-1)/2$ such hyperplanes and by specifying the sense of an inequality associated with each (i.e., $A_{vw}(U) \geq 0$ or $-A_{vw}(U) \geq 0$) we define a convex set (not necessarily bounded) which we may denote by $A_t^* U + B_t^* \geq 0$ where $t = 1,\ldots,2^{V(V-1)/2}$. The intersection of the convex set defined by $A_t^* U + B_t^* \geq 0$ and the convex set defined by $AU + B \overset{>}{(\geq)} 0$ is a bounded convex subset of the latter set and may be defined by the system of linear inequalities

$$\begin{pmatrix} A \\ \cdots \\ A_t^* \end{pmatrix} U + \begin{pmatrix} B \\ \cdots \\ B_t^* \end{pmatrix} \overset{>}{(\geq)} 0$$

$$A_t^{**} \; U + B_t^{**} \overset{>}{(\geq)} 0$$

Some of the $2^{V(V-1)/2}$ subsets may be empty because the system $A_t^* U + B_t^* \geq 0$ is inconsistent or cannot be satisfied by any point in the set defined by $AU + B \overset{>}{(\geq)} 0$.

Each system $A_t^{**} U + B_t^{**} \overset{>}{(\geq)} 0$ restricts U in the same manner as did the system $AU + B \overset{>}{(\geq)} 0$ although more stringently. Therefore we may follow the same procedure for this system as we did for the original system, generating the extreme points of the convex set (some of which may have been determined already

while generating the extreme points of the original convex set), computing the appropriate C matrices, and determining the corresponding optimal strategies. If one strategy is optimal at every extreme point of the subset, it is optimal at every point in the subset. Otherwise we may partition this subset into smaller subsets by the same procedure and examine the extreme points of these sub-subsets etc. By following such a course for $t = 1,\ldots,2^{V(V-1)/2}$ we must eventually partition the convex set defined by $AU + B \overset{>}{(\geq)} 0$ into subsets in each of which a particular strategy is optimal.

In Appendix I we present a method for incorporating the additional constraints given by $A_t^{*}U + B_t^{*} \geq 0$ into the system $AU + B \overset{>}{(\geq)} 0$, and for generating the extreme points of the bounded convex set defined by $A_t^{**}U + B_t^{**} \overset{>}{(\geq)} 0$. In the following chapter we present methods by which this partitioning and evaluation can be carried out without difficulty for the case $N = 1$. When $N = 2$ the analysis can be accomplished without excessive difficulty with the aid of a graph. As N and the number of alternative strategies to be evaluated grow large, the computations involved in a complete analysis could become (although not necessarily) prohibitive. In addition when working in higher dimensions where our geometrical insight is weak and when there are a large number of alternative strategies, we may encounter difficulty in interpreting the implications of a partitioning of the set.

There seem to be essentially two courses open to us when we discover that no single strategy is optimal at all points of the convex set defined by $AU - B \overset{>}{(\geq)} 0$. We may attempt to determine the parameters with greater precision (hopefully sufficient to restrict them within a subset over which one strategy is optimal) or we may simply select one strategy for use. The choice between these courses, the methods to be employed in pursuing the chosen course, and the usefulness of a partitioning of the set in aiding decisions on these matters cannot be determined outside the context of a specific problem. Here we confine ourselves to some examples of when a partitioning could or could not be useful.

If we adopt the alternative of determining the parameters with greater precision, the size, shape, and location of the subsets over which particular strategies are optimal may provide information as to the parameters the additional restriction of which will be most conducive to the specification of a set over which one strategy is optimal and as to the form in which additional restrictions may be sought. On the other hand if the techniques at our disposal in a specific problem do not permit us to be selective in either the parameters to be restricted further or the form of additional restrictions, a partitioning of the set will not be especially useful. This would be the situation when the parameters are coefficients of a polynomial value function bracketed by two polynomials.

If the bracketing polynomials can be determined only by the method of section III A 2c of Chapter 2, we cannot further restrict, say, only the coefficient of the quartic term; we must "squeeze together" the two bracketing polynomials which will usually alter the restrictions on all of the coefficients of the bracketed polynomial.

If we adopt the latter alternative we must apply some additional criterion (perhaps a very informal one) for selecting the strategy to be used. A possible rule is to select this strategy from among those strategies which are optimal for some subset by the criterion that its "optimality subset" be the largest. This would be an application of what is often called the "Bayes Criterion" or the "Principle of Insufficient Reason" over the convex set defined by $AU + B \overset{>}{(=)} 0$. In this case a partitioning of the convex set would be essential to the decision. In some other cases it may be irrelevant. For example if $u_1, \ldots, u_N$ are unit values associated with qualitative objectives, we may use the Churchman-Ackoff procedure (section III A 1 of Chapter 2) to obtain the linear constraints $AU + B \overset{>}{(=)} 0$. In applying this procedure we obtain a specific point in the convex set by manipulating tentative unit values to conform to the linear constraints, and may develop a "feel" for the problem which leads us to believe that in some sense this point is our best estimate of U. We might then decide to use the strategy which is optimal at this particular point even

though it is not optimal at every point in the convex set. Here the specification of the subsets within which various strategies are optimal would not aid in the final decision.

Before leaving the discussion of partitioning we note that when working with a system containing strict linear inequalities some extreme points of the set are not contained in the set. This fact presents no difficulty if we find one strategy to be optimal at every extreme point; that strategy is optimal over the set regardless of whether the inequalities are strict or ordinary. However when no single strategy is optimal over a convex set defined by a particularly unfortunate system of inequalities, it is possible to have a strategy, say s_V, which is optimal at as many as N extreme points of the convex set without being optimal at any point in the set. This occurs when the set over which s_V is optimal intersects the set defined by $AU + B \geqq 0$ but does not intersect the set defined by $AU + B \overset{>}{(\geqq)} 0$. Such a situation would be revealed by the fact that one of the $2^{V(V-1)/2}$ systems $A_t^{**} U + B_t^{**} \overset{>}{(\geqq)} 0$ is inconsistent while the corresponding system of all ordinary inequalities $A_t^{**} U + B_t^{**} \geqq 0$ is consistent. A means of identifying this situation is included in Appendix I.

Before turning to an illustration of the method proposed here, we append a result which in some cases will suffice to specify an optimal strategy over the convex set.

If S_o is given by $tr(S_oC_m) = \min_k tr(S_kC_m)$ for $m = 0,1,\ldots,M$ then s_o is optimal for all $U \geq 0$.

To prove this result we consider a strategy matrix S_k. By hypothesis

$$tr(S_oC_m) \leq tr(S_kC_m) \quad \text{for } m = 0,1,\ldots,M$$

Since $0 \leq u_m$ $\quad u_m\, tr(S_oC_m) \leq u_m\, tr(S_kC_m)$ for $m = 1,\ldots,M$

Summing $\quad tr(S_oC_o) + \sum_{m=1}^{m=M} u_m\, tr(S_oC_m) \leq tr(S_kC_o) + \sum_{m=1}^{m=M} u_m\, tr(S_kC_m)$

$$tr(S_oC) \leq tr(S_kC)$$

Therefore s_o is an optimal strategy for any $U \geq 0$.

I B. Example of Sensitivity Analysis: In order to illustrate the analysis of the sensitivity of the optimal strategy to variations in several parameters, we return to the example given in Chapter 1. In Flagle and Lechat (1963) the interests of four mutually exclusive segments of society were considered to be relevant to the specification of the loss incurred when an action (a_1: hospitalize, a_2: treat as outpatient, a_3: discharge) is taken in the presence of a state of nature (θ_1: infected and contagious, θ_2: infected but non-contagious). These four segments were

1. the productive society as an entity, to be differentiated from the following particular segments

which are directly involved in the diagnostic problem.

2. the health service system as an aggregate of scarce resources
3. the individual patient who is infected with leprosy
4. the health service staff as a professional community

The objectives listed in Chapter 2 (control the disease in the patient and minimize new infections, treatment time, and monetary cost of treatment) and perhaps other objectives, when considered from the points of view of the four segments of society, create different loss matrices such as the following which are presumed to be in the same units. E_m ($m=1,\ldots,4$) is the loss matrix of the segment of society which is m-th in the list above.

$$E_1 = \begin{pmatrix} .153 & .153 \\ .600 & .092 \\ 2.358 & .200 \end{pmatrix} \quad E_2 = \begin{pmatrix} 1.833 & 1.833 \\ .500 & .500 \\ 0 & 0 \end{pmatrix} \quad E_3 = \begin{pmatrix} .084 & .084 \\ .282 & .282 \\ 1.000 & 1.000 \end{pmatrix} \quad E_4 = \begin{pmatrix} 0 & 0 \\ 2.111 & .056 \\ 2.647 & .235 \end{pmatrix}$$

Because $C_m = E_m QP$ each of these loss matrices leads to a different criterion matrix as shown below using the probability distributions of Chapter 1.

$$C_1 = \begin{pmatrix} .012 & .141 \\ .048 & .095 \\ .189 & .227 \end{pmatrix} \quad C_2 = \begin{pmatrix} .147 & 1.686 \\ .040 & .460 \\ 0 & 0 \end{pmatrix} \quad C_3 = \begin{pmatrix} .007 & .077 \\ .023 & .259 \\ .080 & .920 \end{pmatrix} \quad C_4 = \begin{pmatrix} 0 & 0 \\ .170 & .093 \\ .212 & .264 \end{pmatrix}$$

When the interests of the various segments of society are brought to bear on the diagnostic problem individually, they lead to different optimal strategies. From the point of view of productive

society s_2 (hospitalize apparently contagious patients and treat others as outpatient is optimal. If conservation of the health service's resources for other uses were the only consideration, no leper would ever be treated and s_9 (discharge everyone) would be optimal. From the viewpoints of the patient and the staff s_1 (hospitalize every infected person, contagious or not) is optimal. The use of a strategy which is optimal for any one segment of society will impose a penalty on one or more of the other segments, e.g. the use of s_1 will penalize productive society an amount given by $\mathrm{tr}\ (S_1-S_2)C_1 = .056$ while on the other hand the use of s_2 will penalize the patient an amount given by $\mathrm{tr}\ (S_2-S_1)C_3 = .182$. Similar comparisons can be made between other strategies and segments of society.

The individual or group charged with the direction of a health service program must settle on a strategy to be used which represents a suitable compromise of the conflicting interests. This may be done by formally or informally assigning appropriate weights to the interests of the four relevant segments of society. We suppose that in a first attempt to determine a suitable compromise we make only the simplest judgements about the weights to be assigned. We merely rank them and express the weights to be attached to the interests of the health service system (u_2), the patient (u_3), and the staff (u_4) as fractions of the weight attached

to the interests of productive society (u_1) which we set equal to unity. We may summarize one such ranking by the following four inequalities.

$$\begin{array}{c} (A) \\ (B) \\ (C) \\ (D) \end{array} \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \\ u_4 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The four inequalities define a bounded convex set in 3-space and since we can write a loss matrix as the weighted sum

$$L = E_1 + \sum_{m=2}^{m=4} u_m E_m$$

we may use the method of this chapter to determine if such a ranking is sufficient to specify a single optimal strategy and if it is not we may examine the subsets within which particular strategies are optimal to aid in deciding what additional analysis should be carried out. The convex set with which we are working is simple enough so that no formal analysis is needed to find its extreme points which are not themselves in the set because we have a system of strict inequalities. The extreme points are given below with the corresponding C matrices and optimal strategies.

$U' = (1,u_2,u_3,u_4)$	$C = C_1 + \sum_{m=2}^{m=4} u_m C_m$	Optimal Strategies	Constraints Holding as Equalities
(1) (1,0,0,0)	$\begin{pmatrix} \underline{.012} & .141 \\ .048 & \underline{.095} \\ .189 & .227 \end{pmatrix}$	S_2	(B)(C)(D)
(2) (1,1,0,0)	$\begin{pmatrix} .159 & 1.827 \\ \underline{.083} & .555 \\ .189 & \underline{.227} \end{pmatrix}$	S_6	(A)(C)(D)
(3) (1,1,1,0)	$\begin{pmatrix} .166 & 1.905 \\ \underline{.111} & \underline{.815} \\ .269 & 1.147 \end{pmatrix}$	S_5	(A)(B)(D)
(4) (1,1,1,1)	$\begin{pmatrix} \underline{.166} & 1.905 \\ .280 & \underline{.907} \\ .480 & 1.412 \end{pmatrix}$	S_2	(A)(B)(C)

In this problem our ranking of the weights eliminates the two "extreme" strategies s_1 (hospitalize every infected patient) and s_9 (conserve the resources of the health service system for other purposes), but it is not sufficient to eliminate all except one "compromise" strategy. We have at least s_2 (hospitalize apparently contagious cases and treat other infected persons as outpatients), s_5 (treat all infected persons as outpatients), and s_6 (treat only apparently contagious cases and treat them as outpatients) as potential optimal strategies. Having failed to find a strategy which is optimal at every extreme point of the convex set, we may wish to examine the subsets over which particular strategies are optimal.

The first step in partitioning the set is to determine the planes on which the expected losses from the V = 3 strategies found to be optimal at some extreme point of the set are equal. The expected losses for the three strategies are

$$E(s_2) = tr(S_2C) = .107 + .607\ u_2 + .266\ u_3 + .093\ u_4$$

$$E(s_5) = tr(S_5C) = .143 + .500\ u_2 + .282\ u_3 + .262\ u_4$$

$$E(s_6) = tr(S_6C) = .275 + .040\ u_2 + .943\ u_3 + .433\ u_4$$

and the planes on which the expected losses are equal are given by

$$0 = A_{25}(U) = tr((S_2-S_5)C) = -.036 + .107\ u_2 - .016\ u_3 - .169\ u_4$$

$$0 = A_{26}(U) = tr((S_2-S_6)C) = -.168 + .567\ u_2 - .676\ u_3 - .341\ u_4$$

$$0 = A_{56}(U) = tr((S_5-S_6)C) = -.132 + .460\ u_2 - .661\ u_3 - .172\ u_4$$

Since V = 3 there are V(V - 1)/2 = 3 such planes and $2^{V(V - 1)/2}$ = 8 subsets to be examined. These subsets are defined by the four constraints representing the ranking of the weights and the inequalities noted below.

Subset	Constraints			Extreme Points	Optimal Strategies
	A_{25}	A_{56}	A_{26}		
I	≥ 0	≥ 0	≥ 0	2,5,6,7,8,9	S_6
II	≥ 0	≥ 0	≤ 0	5,6	$S_2S_3S_5S_6$
III	≥ 0	≤ 0	≥ 0	5,6,8,10	S_5
IV	≥ 0	≤ 0	≤ 0	3,5,6,10,11,12	S_5

V	≤ 0	≥ 0	≥ 0	5,6,7.9,13,14	S_3
VI	≤ 0	≥ 0	≤ 0	5,6,13,14,15.16	S_3
VII	≤ 0	≤ 0	≥ 0	5,6	$S_2S_3S_5S_6$
VIII	≤ 0	≤ 0	≤ 0	1,4,5,6,11,12,15	S_2

By applying the procedure presented in Appendix I we generate the extreme points of each of these subsets. The designations of the extreme points of each subset are listed above while the extreme points themselves (other than the four given earlier) are tabulated below.

$U' = (1, u_2, u_3, u_4)$	$C = C_1 + \sum_{m=2}^{m=4} u_m C_m$	Optimal Strategies	Constraints Holding as Equalities
(5) (1,.341,.037,0)	$\begin{pmatrix} \underline{.063} & .719 \\ \underline{.063} & \underline{.262} \\ .192 & \underline{.262} \end{pmatrix}$	$S_2S_3S_5S_6$	(D) $A_{25}A_{56}A_{26}$
(6) (1,1,.397,.382)	$\begin{pmatrix} \underline{.162} & 1.858 \\ \underline{.162} & \underline{.693} \\ .301 & \underline{.693} \end{pmatrix}$	$S_2S_3S_5S_6$	(A) $A_{25}A_{56}A_{26}$
(7) (1,1,.383,.383)	$\begin{pmatrix} \underline{.161} & 1.857 \\ \underline{.161} & .690 \\ .300 & \underline{.691} \end{pmatrix}$	S_3S_6	(A)(C) A_{25}
(8) (1,1,.496,0)	$\begin{pmatrix} .162 & 1.866 \\ \underline{.099} & \underline{.683} \\ .228 & \underline{.683} \end{pmatrix}$	S_5S_6	(A)(D) A_{56}
(9) (1,.336,0,0)	$\begin{pmatrix} \underline{.061} & .707 \\ \underline{.061} & .249 \\ .189 & \underline{.227} \end{pmatrix}$	S_3S_6	(C)(D) A_{25}

(10) (1,1,.589,0) $\begin{pmatrix} .163 & 1.873 \\ \underline{.101} & \underline{.708} \\ .236 & .769 \end{pmatrix}$ S_5 (A)(D) A_{26}

(11) (1,1,1,.325) $\begin{pmatrix} \underline{.166} & 1.905 \\ \underline{.166} & \underline{.844} \\ .337 & 1.233 \end{pmatrix}$ S_2S_5 (A)(B) A_{25}

(12) (1,.395,.395,0) $\begin{pmatrix} \underline{.073} & .837 \\ .073 & \underline{.379} \\ .220 & .590 \end{pmatrix}$ S_2S_5 (B)(D) A_{25}

(13) (1,1,.392,.392) $\begin{pmatrix} \underline{.161} & 1.858 \\ .163 & .693 \\ .303 & \underline{.691} \end{pmatrix}$ S_3 (A)(C) A_{26}

(14) (1,.297,0,0) $\begin{pmatrix} \underline{.056} & .642 \\ .060 & .231 \\ .189 & \underline{.227} \end{pmatrix}$ S_3 (C)(D) A_{26}

(15) (1,1,.394,.394) $\begin{pmatrix} \underline{.161} & 1.858 \\ .163 & \underline{.693} \\ .304 & \underline{.693} \end{pmatrix}$ S_2S_3 (A)(C) A_{56}

(16) (1,.288,0,0) $\begin{pmatrix} \underline{.054} & .626 \\ .060 & \underline{.227} \\ .189 & \underline{.227} \end{pmatrix}$ S_2S_3 (C)(D) A_{56}

Upon examining the eight subsets we note that III and IV differ only in the sense of the inequality attached to $A_{26}(U)$. By discarding $A_{26}(U)$ we obtain the union of III and IV, a convex subset in which s_5 is optimal. Similarly by discarding $A_{26}(U)$ we obtain the union of V and VI, a convex subset over which s_3 is optimal. In applying the procedure of Appendix I to the constraints defining subsets I and VIII, we discover that the inequality assigned to $A_{26}(U)$ is redundant in that it is automatically satisfied by every point which satisfies the other constraints. Furthermore we note that subsets II and

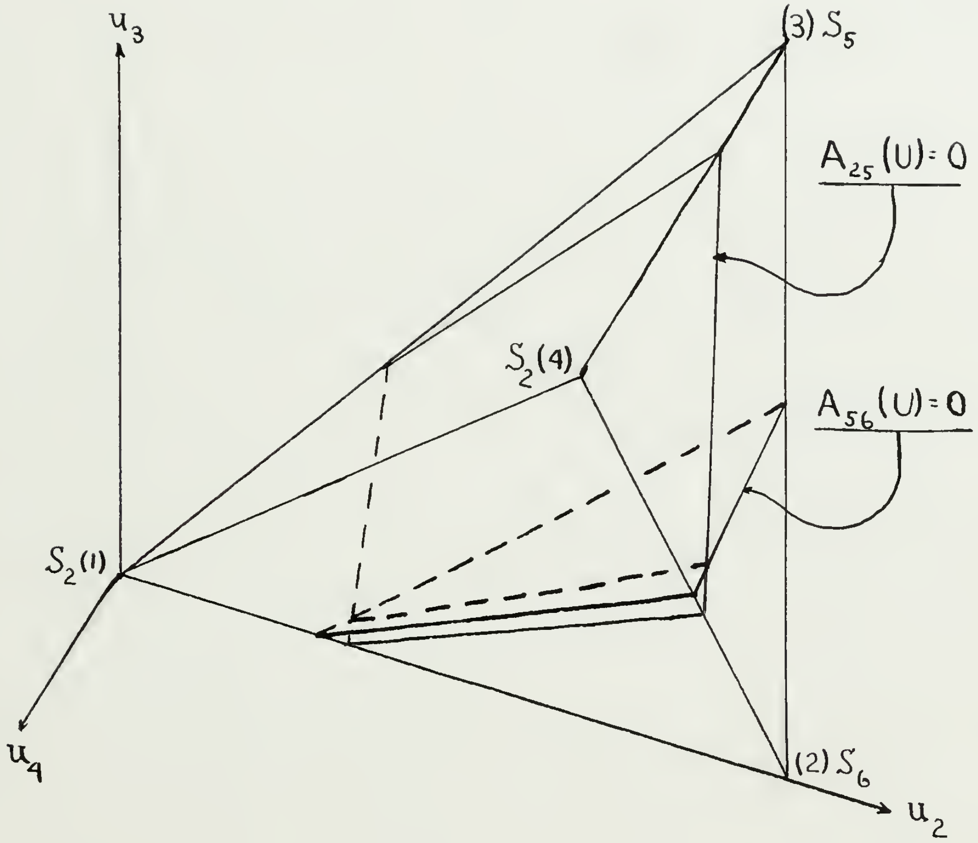
VII are identical; the intersection of the three planes

$$A_{25}(U) = 0$$
$$A_{26}(U) = 0$$
$$A_{56}(U) = 0$$

is a line containing points (5) and (6). Of course the intersection of any two of the three planes is the same line so that one of them is redundant. As a matter of convenience we let $A_{26}(U)$ be the redundant constraint. Thus $A_{26}(U)$ need not be considered in defining any of the subsets.

We have partitioned the original convex set into four convex subsets in each of which a particular strategy is optimal; the union of II and VII is a line which is the intersection of the four convex subsets and belongs to each of them. These subsets are pictured on the following page. The union of III and IV in which s_5 is optimal is the subset nearest the vertex (3). The subset nearest the vertex (2) is I in which s_6 is optimal while s_2 is optimal in the subset nearest the vertices (1) and (4), i.e. in subset VIII. The union of V and VI in which s_3 is optimal is the long narrow subset having as one of its edges the line

$$A_{25}(U) = A_{56}(U) = 0$$



Having completed the mechanistic part of the analysis we move into the realm of art. We must decide whether to select a strategy for use now or to restrict the weights more stringently. We tend to prefer the latter course because intuition suggests that we should be able to say more about the relative weights than is contained in a simple ranking. Furthermore the optimal strategy is relatively insensitive to u_2, u_3, and

u_4 except in the region near the line $A_{25}(U) = A_{56}(U)$ so there is some basis to hope that more precise information on the admissible values of u_2, u_3, and u_4 will restrict them to lie largely within a subset over which a single strategy is optimal.

Suppose that we employ some formal procedure such as the Churchman-Ackoff procedure (Chapter 2) to restrict u_2, u_3, and u_4 more closely perhaps asking questions like, "Should the interests of productive society be weighted more heavily than the interests of the health service staff and the patient taken together?" Such a procedure may yield the following restrictions on u_2, u_3, and u_4.

$$\begin{array}{c} (E) \\ (F) \\ (G) \\ (H) \end{array} \qquad \begin{pmatrix} 1 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \\ u_4 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The constraints derived by ranking the weights are incorporated in these constraints.

The first question to which we address ourselves is whether or not there is a single strategy that is optimal for any compromise satisfying these additional restrictions. In Appendix I (in the context of another problem) we determine the extreme points of the bounded convex set of defined by (E)

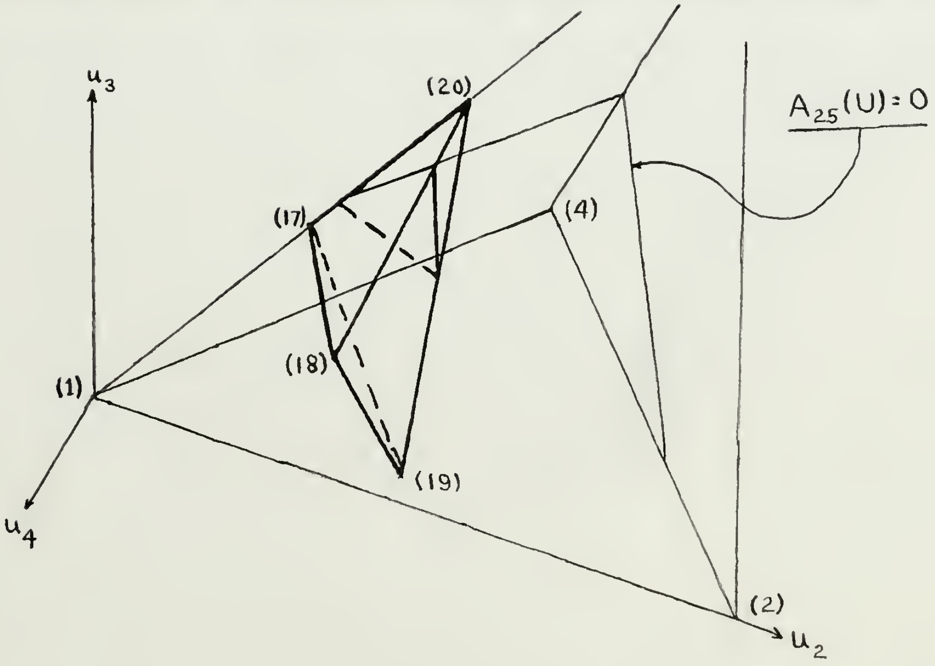
through (H). Given these extreme points we compute the corresponding C matrices and determine the optimal strategies.

$U' = (1, u_2, u_3, u_4)$	$C = C_1 + \sum_{m=2}^{m=4} u_m C_m$	Optimal Strategies	Constraints Holding as Equalities
(17) (1,.333,.333,0)	$\begin{pmatrix} \underline{.063} & .729 \\ .069 & \underline{.325} \\ .215 & .534 \end{pmatrix}$	S_2	(E)(F)(H)
(18) (1,.400,.200,.200)	$\begin{pmatrix} \underline{.072} & .831 \\ .102 & \underline{.349} \\ .247 & .464 \end{pmatrix}$	S_2	(E)(F)(G)
(19) (1,.500,.167,.167)	$\begin{pmatrix} \underline{.082} & .997 \\ .100 & \underline{.383} \\ .237 & .425 \end{pmatrix}$	S_2	(F)(G)(H)
(20) (1,.500,.500,0)	$\begin{pmatrix} .089 & 1.023 \\ \underline{.079} & \underline{.454} \\ .229 & .689 \end{pmatrix}$	S_5	(E)(G)(H)

We still have not restricted u_2, u_3, and u_4 to lie within a set over which a single strategy is optimal. Therefore we must again decide either to select a strategy for use or to attempt to restrict the weights more stringently. In either case a partitioning of this convex set may be helpful. We partition the set by augmenting the inequality $A_{25}(U) \geq 0$ to (E) through (H). These five inequalities define a subset with four extreme points (see Appendix I), (20) and the three given below.

$U' = (1, u_2, u_3, u_4)$	$C = C_1 + \sum_{m=2}^{m=4} u_m C_m$	Optimal Strategies	Constraints Holding as Equalities
(21) (1,.395,.395,0)	$\begin{pmatrix} \underline{.073} & .837 \\ \underline{.073} & \underline{.379} \\ .220 & .590 \end{pmatrix}$	S_2 S_5	(E)(H) A_{25}
(22) (1,.476,.428,.048)	$\begin{pmatrix} \underline{.085} & .976 \\ \underline{.085} & \underline{.429} \\ .233 & .633 \end{pmatrix}$	S_2 S_5	(E)(G) A_{25}
(23) (1,.500,.361,.070)	$\begin{pmatrix} \underline{.088} & 1.012 \\ \underline{.088} & \underline{.425} \\ .232 & .577 \end{pmatrix}$	S_2 S_5	(G)(H) A_{25}

Therefore s_5 is optimal over the subset. If we now augment $A_{25}(U) \leq 0$ to (E) through (H), we define a subset with (17), (18), (19), (21), (22), and (23) as extreme points over which s_2 is optimal. These subsets are pictured below embedded in the convex set obtained by ranking the weights. The subset over which s_5 is optimal is the smaller tetrahedron near the upper-most vertex of the tetrahedron defined by (E) through (H).



If we now attempt to determine the weights with greater precision, the fact that the largest u_4 co-ordinate among (21), (22), and (23) is .07 suggests that we can restrict the admissible values of u_2, u_3, and u_4 to lie within a set over which s_2 is optimal if we can satisfy ourselves that u_4 should be greater than .07 in addition to satisfying (E) through (H). Of course other supplementary restrictions can achieve the same end, e.g. a restriction of u_3 to be less than .7 u_2.

Alternately we may select a strategy for use without attempting to further restrict u_2, u_3, and u_4. If we decide to use the strategy having the largest "optimality subset," we will select s_2 since the volume of the subset in which it is optimal is larger than the volume of the subset in which s_5 is optimal. Perhaps in the course of establishing the constraints on u_2, u_3, and u_4 summarized by (E) through (H), we receive the impression that (1,.450,.267,.142) is a "good" estimate of the proper weights to be given to the interests of the four relevant segments of society. If we rely on this estimate and use the relation

$$L = E_1 + \sum_{m=2}^{m=4} u_m E_m$$

we obtain the loss matrix used for this problem in Chapter 1, and as was shown there s_2 is optimal. The same conclusion follows from the fact that the point lies within the subset over which s_2 is optimal.

Regardless of the strategy employed or the criterion by which it is selected, it is of some interest to know what range of expected losses may be encountered using this strategy. If we elect to use s_2 we can determine this range from the maximum and minimum values oftained by computing

$$E\ (s_2) = tr\ (S_2C) = .107 + .607\ u_2 + .266\ u_3 + .093\ u_4$$

at the extreme points of the convex set within which U is constrained. This equation for $E(s_2)$ need not be used explicitly since we can obtain the desired values more expeditiously by computing $tr(S_2C)$ for the C matrices corresponding to the extreme points. Carrying out this computation at the four extreme points of the convex set defined by (E) through (H) we find

Extreme Point	(17)	(18)	(19)	(20)
$E(s_2)$	.398	.421	.470	.543

Therefore the range of expected losses that may be encountered using s_2 is

$$.398 < E(s_2) < .543$$

Before leaving this example we again point out the linearity assumptions of this method. It is possible in such a problem that we would also want to investigate the sensitivity of the optimal strategy to variations in other parameters, e.g. some of the entries in the loss matrices of the four segments of

society. Because $u_1 = 1$ we may use the method of this chapter to analyze u_2, u_3, and u_4 concurrent with an analysis of variations in any of the elements of E_1 (e.g. the loss incurred by failing to treat a noncontagious patient) as long as these losses are subject only to linear constraints. We cannot however analyze for variations in the elements of E_2, E_3, or E_4 in an analysis including u_2, u_3, or u_4 respectively because we could not then express the loss matrix as a linear function of the parameters under analysis. Of course we do not rule out "common sense" analysis. For example given the restrictions (E) through (H) on U, any increase in the elements associated with a_3 in any loss matrix will not alter the optimal strategy; we could replace the ones in the last row of E_3 by 10,000 or more without influencing the optimal strategy.

II. Specification of Optimality Region for a Given Strategy:

In some applications it may be more natural to view sensitivity analysis in a manner other than the one we have presented thus far in this chapter. Instead of examining those strategies which are optimal for various U in a given convex set, we may wish to specify the convex set in which U must lie for a given strategy to be optimal. This section is devoted to a discussion of a method for specifying such a set when the M parameters under scrutiny, $U' = (u_1, \ldots, u_M)$, satisfy condition 1, i.e. the criterion matrix can be expressed as a linear function

$$C = C_o + \sum_{m=1}^{m=M} u_m C_m$$

where C_m is an arbitrary JxI matrix for $m = 0,1,\ldots,M$.

II A. General Discussion: We consider a strategy, s_o, having a given strategy matrix, S_o, and we adopt the notational convention that the one in row i of S_o appears in column j_i, i.e. strategy s_o dictates that we take action a_{j_i} when indication x_i is observed. It follows immediately from our definition of a criterion matrix that s_o is an optimal strategy for any criterion matrix such that

$$c_i = \underset{j}{\mathrm{Min}}\, c_{ji} = c_{j_i i} \quad \text{for } i = 1,\ldots,I$$

or equivalently s_o is optimal for any criterion matrix in which the following I(J-1) inequalities hold

$$c_{ji} - c_{j_i i} \geq 0 \quad j_i \neq j = 1,\ldots,J \quad \text{and } i = 1,\ldots,I$$

Because condition 1 holds for the elements of U, each entry in C is a linear function of the elements of U, and the difference between any two entries in C is also a linear function of the elements of U. Therefore the I(J-1) constraints specifying the optimality of s_o(together with any prior linear restrictions on U) form a system of linear constraints on U which can be expressed as

$$AU + B \overset{>}{(\geq)} 0$$

presuming again that any linear equality constraints on U have been used to reduce the dimensionality of U. Therefore condition 2. for the applicability of the methods of sensitivity analysis will be satisfied.

Having demonstrated that we can specify by a system of linear constraints the set over which a given strategy, s_o, is optimal, we turn to the question of the usefulness of our results. No definitive answer to such a question can be given outside the context of a specific problem because the usefulness of the specification is predicated on our ability to interpret the system of linear constraints in terms of the parameters of which we are dubious. However we can suggest that an imaginative use of the methods for investigating the properties of a convex set defined by linear constraints presented in Appendix I may permit us to distill from the abstract system of constraints results which have intuitive meaning in the context of the problem at hand.

The first result of an application of the methods of Appendix I will be either the determination of one U (in particular an extreme point of the convex set) for which s_o is optimal or the discovery that the system $AU + B \underset{(=)}{>} 0$ is inconsistent. The latter result may well occur if we arbitrarily select a strategy and attempt to define a set over which it is optimal; there may be no admissible value of U for which the strategy is optimal. We may also determine if the convex set

over which s_o is optimal is bounded or unbounded. This knowledge may be of interest in itself and/or helpful in interpreting the results of the methods discussed below.

Appendix I includes a method for testing each constraint in the system $AU + B \underset{(=)}{>} 0$ for redundancy. The bulk of the system (the number I(J-1) often will be large) may be a major obstruction in the way of understanding the implications of the system of constraints. However at least some of these I(J-1) constraints usually will be redundant. Therefore the redundancy tests of Appendix I sometimes can reduce the size of the system of constraints which we must attempt to understand.

Finally in Appendix I we discuss how the extreme points of the convex set defined by $AU + B \underset{(=)}{>} 0$ may be computed. In some cases (particularly those in which the constraints have many irrational coefficients) a list of the extreme points of the set will specify the set in a manner which is more readily interpretable than is the system of constraints. However we must bear in mind that if the set is unbounded a list of extreme points is not sufficient to define the convex set.

II B. Example of Sensitivity Analysis: To illustrate the procedure for specifying the set over which a particular strategy is optimal, we return to the example of section I B. There we considered the estimation of weights to be assigned to the interests of four segments of society concerned with

leprosy diagnosis and treatment; here we simply accept the loss matrix

$$L = \begin{pmatrix} 1.0 & 1.0 \\ 1.2 & .4 \\ 3.0 & .5 \end{pmatrix}$$

for purposes of deciding upon an optimal strategy. As we noted in section I B acceptance of this loss matrix is equivalent to the assignment of the weights U' = (1, .450, .267, .142) to the interests of the four segments of society. The loss matrix above is the loss matrix used in Chapter 1 where s_2 with strategy matrix

$$S_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

was found to be optimal. This strategy is to hospitalize those persons who exhibit positive (contagious) indications and to treat as outpatients all other persons having manifest symptoms of leprosy.

Having derived this optimal strategy a natural question is, "If someone were to disagree with the weights which we have imputed to the four segments of society, would they necessarily disagree with our choice of an optimal strategy?" As stated the question is not amenable to a quantitative answer, but we can convert it to a suitable form by asking, "For what values of U will s_2 be an optimal strategy?"

Before attempting to answer this question we set $u_1 = 1$ and require $u_m \geq 0$ for m = 2, 3, 4 in order to eliminate the possibility of a negative weight being assigned. Furthermore we assume a consensus to the effect that the interests of any one of the specific segments of society (health service resources, patient, or professional staff) should not be weighted more heavily than the interests of society as a whole (apart from the three segments considered explicitly), i.e. $u_m \leq 1$ for m = 2, 3, 4. Therefore we have the following six prior constraints which define admissible values of U

$$\begin{array}{c} (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \\ u_4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We next determine the linear constraints that define the set over which s_2 is optimal. These constraints are obtained from the criterion matrix

$$C = C_1 + \sum_{m=2}^{m=4} u_m C_m$$

where the four 3x2 matrices $C_m (m=1,\ldots,4)$ are given in section I B. The $I(J-1) = 2(2) = 4$ constraints that define the set

over which s_2 is optimal are given below.

$$\begin{array}{ll} (7) & c_{12}-c_{22} \geq 0 \\ (8) & c_{32}-c_{22} \geq 0 \\ (9) & c_{21}-c_{11} \geq 0 \\ (10) & c_{31}-c_{11} \geq 0 \end{array} \quad \begin{pmatrix} 1.226 & -.182 & -.093 \\ -.460 & .661 & .172 \\ -.107 & .016 & .169 \\ -.147 & .073 & .212 \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \\ u_4 \end{pmatrix} + \begin{pmatrix} .046 \\ .132 \\ .036 \\ .165 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

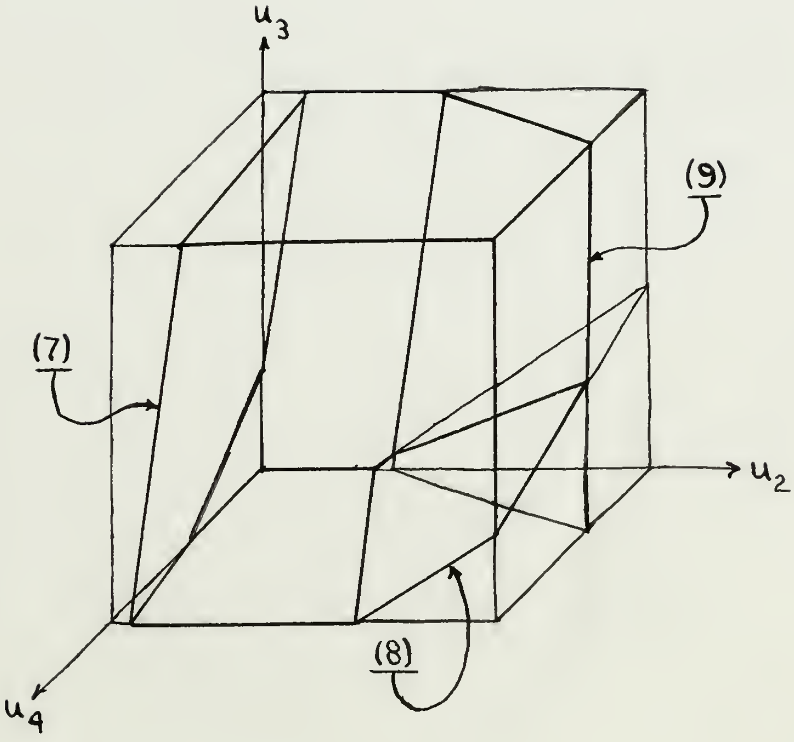
These four inequalities when coupled to the six inequalities given earlier define the set of admissible values of U over which s_2 is optimal.

It is not surprising that an application of the methods of Appendix I reveals the system of ten constraints to be consistent (we started the analysis with a U satisfying all constraints) and shows the set to be bounded (we have $0 \leq u_m \leq 1$ for $m = 2, 3, 4$). We also find that constraint (10) is redundant and therefore we may discard it. The extreme points of the convex set are listed below. Because the system of constraints contains only ordinary inequalities, all extreme points are in the set.

Extreme Points $U' = (u_1, u_2, u_3, u_4)$	Constraints Holding as Equalities
(1 0 0 0)	(1) (2) (3)
(1 1 1 1)	(4) (5) (6)
(1 .288 0 0)	(2) (3) (8)
(1 .458 1 0)	(3) (6) (9)
(1 1 1 .325)	(5) (6) (9)

(1	1	.397	.382)	(5)	(8)	(9)
(1	.341	.037	0)	(3)	(8)	(9)
(1	1	.240	1)	(4)	(6)	(8)
(1	.663	0	1)	(2)	(6)	(8)
(1	.037	0	1)	(2)	(6)	(7)
(1	0	0	.497)	(1)	(2)	(7)
(1	0	.254	0)	(1)	(3)	(7)
(1	.108	1	0)	(3)	(5)	(7)
(1	.181	1	1)	(5)	(6)	(7)

The convex set over which s_2 is optimal is pictured below embedded in the cube which represents the set of all admissible values of U.



The description of the set of admissible values of U over which s_2 is optimal in terms of either the nine constraints or the fourteen extreme points will permit us to say if s_2 is optimal for any given U. However to summarize the situation in our own minds it may be helpful to sacrifice some accuracy by saying that s_2 will be regarded as optimal by everyone except people who assign a large weight to one segment while assigning small weights to the other two segments and people who assign large weights to any two segments while assigning a small weight to the remaining segment. In the region surrounding the principal diagonal of the cube (the line between (1000) and (1111) on which $u_2 = u_3 = u_4$), the status of s_2 as an optimal strategy is relatively insensitive to the particular weights assigned to the interests of the three specific segments of society.

Chapter 4

Sensitivity Analysis for Variations in a Single Parameter

In this chapter we are concerned with one dimensional sensitivity analysis. The assumptions basic to the methods discussed here are the same as those presented earlier but to simplify the exposition we present them in a somewhat different form. The methods of sensitivity analysis for variations in a single scalar parameter u are applicable when

(i.) the criterion matrix can be expressed as the following linear function of the parameter.

$$C^* = C + u\,D$$

where C and D are arbitrary JxI matrices. This condition could arise, for example, if the loss matrix was of the form

$$L^* = L + u\,E$$

where L and E are arbitrary JxN matrices. In such a case we would have

$$C = L/P \quad \text{i.e. } c_{ji} = \sum_{n=1}^{n=N} l_{jn} q_n p_{ni} \qquad j = 1,\ldots,J\ ;\ i = 1,\ldots,I$$

$$D = E/P \quad \text{i.e. } d_{ji} = \sum_{n=1}^{n=N} e_{jn} q_n p_{ni} \qquad j = 1,\ldots,J\ ;\ i = 1,\ldots,I$$

In the following sections we assume that C and D have been computed.

(2.) The parameter under analysis is at most subject to linear inequality constraints of the form

$$k_1 \overset{<}{(\leq)} u \overset{<}{(\leq)} k_2$$

We will assume for convenience of discussion that we start with $u = 0$ and wish to analyze the changes which take place as u increases and/or decreases, that is we will analyze for deviations of the parameter away from some initial value.

The methods of analysis discussed in Chapter 3 are applicable to this one dimensional problem. When the parameter under examination is subject to constraints of the form given in (2) above the methods of analysis discussed in section I of Chapter 3 are applicable. We may set $u = k_1$ and $u = k_2$, and determine the optimal strategies. If a strategy is optimal at both points it is optimal over the interval. If no strategy is optimal at both points we may partition the interval into sub-intervals over which particular strategies are optimal. However if there is not one strategy which is optimal over the entire interval this approach usually will entail a greater number of computations than will the methods presented in this chapter. We may use the method of analysis discussed in section II of Chapter 3 to specify the interval (if any) over which a particular strategy is optimal, but if we begin the analysis with a u for which the strategy in question is optimal, the desired

interval usually can be specified more easily using the methods presented in this chapter.

I. Analysis When D is Arbitrary: In this section we assume that D is an arbitrary JxI matrix while in section II we will make a special assumption about the nature of D which can be used to ease the computational effort involved in sensitivity analysis.

I A. Preliminary Discussion: The second step in the solution procedure outlined in Chapter 1 consists of determining the minimum element of each column of the C matrix. Because of the way we have structured the problem in this chapter, this step consists of determining (given a specific value of u)

$$c_i^* = \min_j c_{ji}^* = \min_j (c_{ji} + u d_{ji}) \quad \text{for } i = 1,\ldots,I$$

We generally expect that c_i^* will be a piecewise linear function of u. As u increases or decreases the subscript j associated with c_i^* will change, and the values of u at which these changes take place are discontinuity points in the slope of c_i^* as a function of u. We denote the values of u at which the subscript j associated with c_i^* changes by

$$u_{1i}, u_{2i}, \ldots, u_{ti}, \ldots \qquad \text{for increasing } u$$

$$u_{-1i}, u_{-2i}, \ldots, u_{-ti}, \ldots \qquad \text{for decreasing } u$$

We denote the subscript j associated with c_i^* by

$$j_o \quad (\text{i.e.} \quad c_i^* = c_{j_o i}^* = \underset{j}{\text{Min}}\, c_{ji}^*) \quad \text{for} \quad u_{-1i} \leq u \leq u_{1i}$$

in particular we have for $u = 0$, $c_i = c_{j_o i} = \underset{j}{\text{Min}}\, c_{ji}$

$$j_1 \quad (\text{i.e.} \quad c_i^* = c_{j_1 i}^* = \underset{j}{\text{Min}}\, c_{ji}^*) \quad \text{for} \quad u_{1i} \leq u \leq u_{2i}$$

$$\vdots$$

$$j_t \quad (\text{i.e.} \quad c_i^* = c_{j_t i}^* = \underset{j}{\text{Min}}\, c_{ji}^*) \quad \text{for} \quad u_{ti} \leq u \leq u_{t+1i}$$

$$\vdots$$

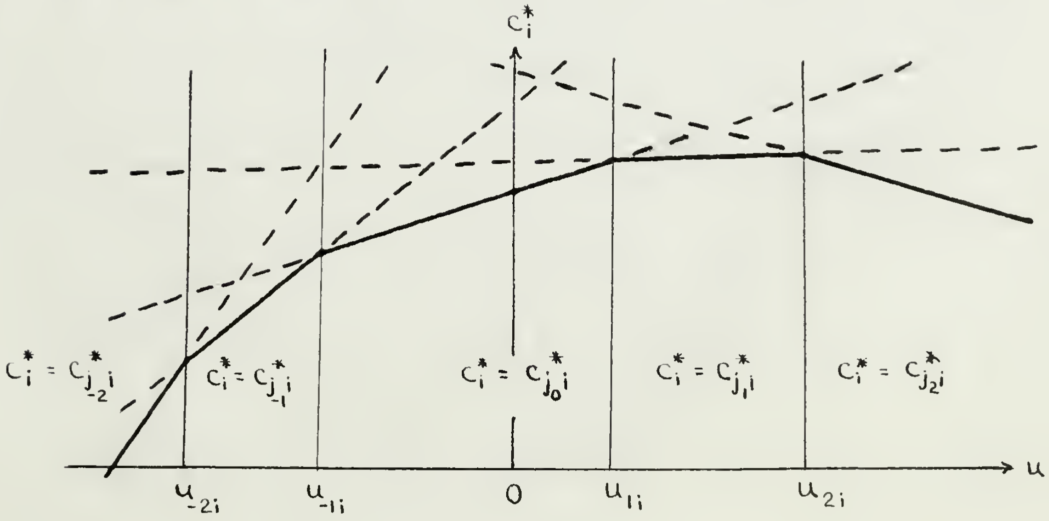
$$j_{-1} \quad (\text{i.e.} \quad c_i^* = c_{j_{-1} i}^* = \underset{j}{\text{Min}}\, c_{ji}^*) \quad \text{for} \quad u_{-2i} \leq u \leq u_{-1i}$$

$$\vdots$$

$$j_{-t} \quad (\text{i.e.} \quad c_i^* = c_{j_{-t} i} = \underset{j}{\text{Min}}\, c_{ji}^*) \quad \text{for} \quad u_{-t-1i} \leq u \leq u_{-ti}$$

$$\vdots$$

The following graph of some c_i^* as a function of u will perhaps help to put this notation in perspective. We note in connection with this graph that c_i^* need not be a monotonic function of u but that it will be concave.



With this notation we observe that u_{1i} is given by

$$u_{1i} = \frac{c_{j_1 i} - c_{j_0 i}}{d_{j_0 i} - d_{j_1 i}} = \operatorname*{Min}_{j} \left(\frac{c_{ji} - c_{j_0 i}}{d_{j_0 i} - d_{ji}} \;\middle|\; d_{ji} < d_{j_0 i} \right)$$

$$\text{and } 0 \leq u_{1i}$$

By hypothesis for $u = 0$, $c_{j_0 i} \leq c_{ji}$ for all j. As u increases only those c^*_{ji} having a slope less than $d_{j_0 i}$ will intersect $c^*_{j_0 i}$. The minimum operator determines the first u for which such an intersection occurs. For $0 \leq u \leq u_{1i}$, $c^*_{j_0 i} \leq c^*_{ji}$ for all $j \neq j_0$; for $u = u_{1i}$, $c^*_{j_0 i} = c^*_{j_1 i} \leq c^*_{ji}$

for all $j \neq j_0, j_1$ and for $u_{1i} < u$, $c^*_{j_1 i} < c^*_{j_0 i}$.

Similarly u_{2i} is given by

$$u_{2i} = \frac{c_{j_2 i} - c_{j_1 i}}{d_{j_1 i} - d_{j_2 i}} = \operatorname*{Min}_{j} \left(\frac{c_{ji} - c_{j_1 i}}{d_{j_1 i} - d_{ji}} \;\middle|\; d_{ji} < d_{j_1 i} \right)$$

$$\text{and } u_{1i} \leq u_{2i}$$

This point is determined by the intersection of the lines $c^*_{j_1 i}$ and $c^*_{j_2 i}$. We prove that $u_{1i} \leq u_{2i}$ as a by-product of the proof that

$$u_{1i} \leq \frac{c_{j_2 i} - c_{j_0 i}}{d_{j_0 i} - d_{j_2 i}} \leq u_{2i}$$

The requirements $d_{j_2 i} < d_{j_1 i} < d_{j_0 i}$ insure that $0 < (d_{j_0 i} - d_{j_2 i})$. For $u = u_{1i}$ we have $c^*_{j_0 i} = \underset{j}{\mathrm{Min}}\, c^*_{ji}$ which implies

$$c_{j_0 i} + u_{1i}\, d_{j_0 i} \leq c_{j_2 i} + u_{1i}\, d_{j_2 i}$$

$$u_{1i}\,(d_{j_0 i} - d_{j_2 i}) \leq c_{j_2 i} - c_{j_0 i}$$

For $u = u_{2i}$ we have $c^*_{j_2 i} = \underset{j}{\mathrm{Min}}\, c^*_{ji}$ which implies

$$c_{j_2 i} + u_{2i}\, d_{j_2 i} \leq c_{j_0 i} + u_{2i}\, d_{j_0 i}$$

$$c_{j_2 i} - c_{j_0 i} \leq u_{2i}\,(d_{j_0 i} - d_{j_2 i})$$

Taken together these two results are

$$u_{1i}\,(d_{j_0 i} - d_{j_2 i}) \leq c_{j_2 i} - c_{j_0 i} \leq u_{2i}\,(d_{j_0 i} - d_{j_2 i})$$

Division of the three terms by $(d_{j_0 i} - d_{j_2 i})$ will not alter the senses of the inequalities and will produce the desired result. The general relation from which u_{1i} and u_{2i} may be obtained by setting $t = 0$ and 1 respectively is

$$u_{t+1 i} = \frac{c_{j_{t+1} i} - c_{j_t i}}{d_{j_t i} - d_{j_{t+1} i}} = \underset{j}{\mathrm{Min}} \left(\frac{c_{ji} - c_{j_t i}}{d_{j_t i} - d_{ji}} \,\middle|\, d_{ji} < d_{j_t i} \right)$$

The general proof that

$$u_{ti} \leq \frac{c_{j_{t+1}i} - c_{j_{t-1}i}}{d_{j_{t-1}i} - d_{j_{t+1}i}} \leq u_{t+1\ i}$$

is the same as the proof given above for the special case t = 1.

With a finite number of terminal actions the repeated calculation of u_{ti} must eventually bring us to a stage, T, where $d_{j_T i} \leq d_{ji}$ for all j. Then the criterion for $u_{T+1\ i}$ cannot be applied and we have $c^*_i = c^*_{j_T i}$ for all u such that $u_{Ti} \leq u$. We may formally describe this situation by stating that $u_{T+1\ i} = \infty$.

The relations for decreasing u are mirror images of the ones described for increasing u. We have u_{-1i} given by

$$u_{-1i} = \frac{c_{j_{-1}i} - c_{j_0 i}}{d_{j_0 i} - d_{j_{-1}i}} = \underset{j}{\mathrm{Max}} \left(\frac{c_{ji} - c_{j_0 i}}{d_{j_0 i} - d_{ji}} \ \middle| \ d_{j_0 i} < d_{ji} \right)$$

$$\text{and } u_{-1i} \leq 0$$

At u = 0, $c_{j_0 i} \leq c_{ji}$ for all j. As u decreases only c^*_{ji} having slopes greater than $d_{j_0 i}$ will intersect $c^*_{j_0 i}$. The maximum operator determines the first u for which such an intersection occurs. For $u_{-1i} \leq u \leq 0$, $c^*_{j_0 i} \leq c^*_{ji}$ for all $j \neq j_0$; for $u_{-1i} = u$, $c^*_{j_0 i} = c^*_{j_{-1}i} \leq c^*_{ji}$ for all $j \neq j_0, j_{-1}$; and for $u < u_{-1i}$, $c^*_{j_{-1}i} < c^*_{j_0 i}$.

Similarly u_{-2i} is determined by

$$u_{-2i} = \frac{c_{j_{-2}i} - c_{j_{-1}i}}{d_{j_{-1}i} - d_{j_{-2}i}} = \operatorname*{Max}_{j} \left(\frac{c_{ji} - c_{j_{-1}i}}{d_{j_{-1}i} - d_{ji}} \;\middle|\; d_{j_{-1}i} < d_{ji} \right)$$

$$\text{and } u_{-2i} \leq u_{-1i}$$

This point is determined by the intersection of the lines $c^*_{j_{-1}i}$ and $c^*_{j_{-2}i}$. The proof that $u_{-2i} \leq u_{-1i}$ follows from the proof that

$$u_{-2i} \leq \frac{c_{j_{-2}i} - c_{j_0 i}}{d_{j_0 i} - d_{j_{-2}i}} \leq u_{-1i}$$

The requirements $d_{j_0 i} < d_{j_{-1}i} < d_{j_{-2}i}$ insure that $(d_{j_0 i} - d_{j_{-2}i}) < 0$. For $u = u_{-1i}$ we have $c^*_{j_0 i} = \operatorname*{Min}_{j} c^*_{ji}$ which implies

$$c_{j_0 i} + u_{-1i}\, d_{j_0 i} \leq c_{j_{-2}i} + u_{-1i}\, d_{j_{-2}i}$$

$$u_{-1i}\,(d_{j_0 i} - d_{j_{-2}i}) \leq c_{j_{-2}i} - c_{j_0 i}$$

For $u = u_{-2i}$ we have $c^*_{j_{-2}i} = \operatorname*{Min}_{j} c^*_{ji}$ which implies

$$c_{j_{-2}i} + u_{-2i}\, d_{j_{-2}i} \leq c_{j_0 i} + u_{-2i}\, d_{j_0 i}$$

$$c_{j_{-2}i} - c_{j_0 i} \leq u_{-2i}\,(d_{j_0 i} - d_{j_{-2}i})$$

Taken together these two results are

$$u_{-1i}\,(d_{j_0 i} - d_{j_{-2} i}) \leq c_{j_{-2} i} - c_{j_0 i} \leq u_{-2i}\,(d_{j_0 i} - d_{j_{-2} i})$$

Division of the three terms by $(d_{j_0 i} - d_{j_{-2} i})$ will reverse the senses of the inequalities and will produce the desired result. The general relation of which u_{-1i} and u_{-2i} are special cases obtained by setting $-t = 0$ and -1 respectively is

$$u_{-t-1i} = \frac{c_{j_{-t-1} i} - c_{j_{-t} i}}{d_{j_{-t} i} - d_{j_{-t-1} i}} = \operatorname*{Max}_{j} \left(\frac{c_{ji} - c_{j_{-t} i}}{d_{j_{-t} i} - d_{ji}} \;\middle|\; d_{j_{-t} i} < d_{ji} \right)$$

The general proof that

$$u_{-t-1i} \leq \frac{c_{j_{-t-1} i} - c_{j_{-t+1} i}}{d_{j_{-t+1} i} - d_{j_{-t-1} i}} \leq u_{-ti}$$

is the same as the proof given above for the special case $-t = -1$.

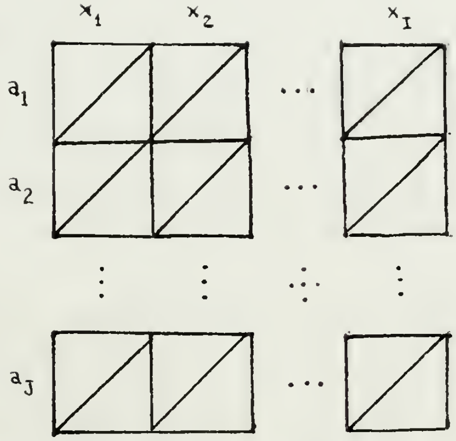
With a finite number of terminal actions the repeated calculation of u_{-ti} must eventually bring us to a stage, say $-T$ (not related to the T for increasing u), where $d_{ji} \leq d_{j_{-T} i}$ for all j. Then the criterion for u_{-T-1i} cannot be applied and we have $c_i^* = c_{j_{-T} i}^*$ for all $u \leq u_{-Ti}$. We may formally describe this situation by stating that $u_{-T-1i} = -\infty$.

In the preceding discussion we have assumed implicitly that the subscripts j_t and j_{-t} are uniquely determined. It is of course possible that there exist cells (ji) and $(j'i)$ such

that $c_{ji} = c_{j'i}$ and $d_{ji} = d_{j'i}$. Thus it is possible that j_t or j_{-t} will not be uniquely determined at some stage. If such an indetermination appears we may take any of the eligible candidates to be j_t or j_{-t}.

I B. Analysis Procedure: In this section we present a step-by-step account of a computational procedure based on the results of the preceding section which will permit us to specify an optimal strategy for any admissible value of u. Assuming that the matrices C and D have been computed, the analysis proceeds as follows.

1. We first construct "Table 1," a JxI table with rows corresponding to $a_1, \ldots, a_J$ and columns corresponding to $x_1, \ldots, x_I$. Each cell of the table may contain two entries and to separate them we cut each cell by a diagonal line as shown below



The rest of the procedure involves the computation of the number (if any) to be entered in each of the 2(IJ) positions in this table. Often the procedure will result in no entry being made in a particular position.

2. Presumably for each column of C we have identified in some manner such as underscoring the minimum element

$$c_i = c_{j_o i} = \operatorname*{Min}_{j} c_{ji}$$

By subtracting c_i from every element in column i of the C matrix, we form a table with entries $c_{ji} - c_{j_o i} = 0$. We emphasize that the subscripts j_o, j_1, j_{-1}, etc. depend on i so that, for example, the cell $(j_o i)$ will be in various rows as we move from column to column.

3. We form a table from D by subtracting from every element of column i the element $d_{j_o i}$ corresponding to $c_{j_o i}$ in the C matrix and by multiplying every element so formed by -1. Thus we have a table containing as its entries $d_{j_o i} - d_{ji}$.

4. Analysis for increasing u

4a. From the result of 3. we form a JxI table with entries

$$\begin{array}{ll} d_{j_o i} - d_{ji} & \text{if } 0 < d_{j_o i} - d_{ji} \\ 0 \text{ (or a blank)} & \text{if } d_{j_o i} - d_{ji} \leq 0 \end{array}$$

4b. With the results of 2. and 4a. we form a JxI table with entries

$$\begin{array}{ll} (c_{ji} - c_{j_o i}) \,/\, (d_{j_o i} - d_{ji}) & \text{if } 0 < d_{j_o i} - d_{ji} \\ \infty \text{ (or a blank)} & \text{if } d_{j_o i} - d_{ji} \leq 0 \end{array}$$

This is done by dividing each entry in the table of 2. by the corresponding entry in the table of 4a., adopting the convention that any number divided by zero is infinity.

4c. For each column in the table of 4b. we mark the minimum entry. This entry is u_{1i} and we denote the cell in which it appears by $(j_1 i)$. We then enter this number in the upper left part of cell $(j_0 i)$ and the lower right of cell $(j_1 i)$ for each column i in Table 1.

4d(i). We form a new table from the result of 4a. by subtracting from each entry in column i the entry in cell $(j_1 i)$. If the remainder is positive enter it in cell (ji) of the new table. Otherwise enter zero or leave the cell blank. In this way we form a table having as its entries

$$d_{j_1 i} - d_{ji} = (d_{j_0 i} - d_{ji}) - (d_{j_0 i} - d_{j_1 i}) \quad \text{if } 0 < d_{j_1 i} - d_{ji}$$

$$0 \text{ (or a blank)} \qquad \text{if } d_{j_1 i} - d_{ji} \leq 0$$

4d(ii). Form a new table from the result of 2. by subtracting from each entry in column i the entry in cell $(j_1 i)$. We thus form a table with entries

$$c_{ji} - c_{j_1 i} = (c_{ji} - c_{j_0 i}) - (c_{j_1 i} - c_{j_0 i})$$

As will be seen in (iii) below this operation need not be executed for those cells (ji) for which the corresponding entries in the result of (i) are nonpositive.

4d(iii). From the results of 4d(i) and 4d(ii) we form a table containing the entries

$$(c_{ji} - c_{j_1 i})/(d_{j_1 i} - d_{ji}) \quad \text{if } 0 < d_{j_1 i} - d_{ji}$$

$$\infty \text{ (or a blank)} \quad \text{if } d_{j_1 i} - d_{ji} \leq 0$$

This is accomplished by dividing each entry in the table of 4d(ii) by the corresponding entry in the table of 4d(i), adopting the convention that any number divided by zero is infinity.

4d(iv). We mark the minimum entry in each column of the table of 4d(iii). This number is u_{2i} and the cell in which it appears is denoted by $(j_2 i)$. We then enter this number in the upper left part of cell $(j_1 i)$ and the lower right part of cell $(j_2 i)$ for each column i in Table 1. Note that we now have specified in cell $(j_1 i)$ the range of u in which $c_i^* = c_{j_1 i}^*$ for $i = 1,\ldots,I$.

4d(v). We then begin a new cycle starting with the subtraction of the number in cell $(j_2 i)$ from every entry in column i for $i = 1,\ldots,I$ in the tables of 4d(i) and 4d(ii). The procedure is continued until at some cycle, t, we find one of the following conditions holding for some column i.

4d(va). If we have a preassigned upper limit k_2 on u we may find that $k_2 \leq u_{ti}$ indicating that the change in the subscript associated with c_i^* from j_{t-1} to j_t occurs outside (or on the

boundary) of the interval within which u must lie. In this case we enter k_2 rather than u_{ti} in the upper left part of cell $(j_{t-1}i)$ in Table 1.

4d(vb). It is possible at some cycle before condition (a) appears for us to encounter the condition

$$u_{ti} = \frac{c_{j_t i} - c_{j_{t-1} i}}{d_{j_{t-1} i} - d_{j_t i}} < k_2 = \underset{j \neq j_t}{\mathrm{Min}} \left(\frac{c_{ji} - c_{j_{t-1} i}}{d_{j_{t-1} i} - d_{ji}} \middle| \; d_{ji} < d_{j_{t-1} i} \right)$$

The change in the subscript associated with c_i^* from j_{t-1} to j_t occurs at an admissible value of u, but the fact that the second smallest entry in column i is not exceeded by k_2 implies that if $u_{t+1\ i}$ is computed it will be found to be outside (or on the border) of the interval within which u must lie. This follows from the statement in Section IA that

$$\frac{c_{j_{t+1} i} - c_{j_{t-1} i}}{d_{j_{t-1} i} - d_{j_{t+1} i}} \leq u_{t+1 i}$$

In this case we enter k_2 rather than u_{t+1i} in the upper left part of cell $(j_t i)$ in Table 1.

4d(vc). Whether or not we have an upper limit on u a sufficiently large number of cycles will bring us to a cycle such that the subtraction of the number in cell $(j_t i)$ will drive every entry non-positive in column i of a table such as the

one in 4d(i) above. This indicates that $d_{j_t i} \leq d_{ji}$ for all j. In this case we may enter ∞ (or k_2 if such an upper limit is specified) in the upper left part of cell $(j_t i)$ in Table 1 to indicate that $c_i^* = c_{j_t i}^*$ for all admissible $u \geq u_{ti}$.

When any one of the conditions holds for a column no further analysis need be carried out on that column and we may simply leave out the corresponding columns in the tables for $(d_{j_t i} - d_{ji})$ and $(c_{ji} - c_{j_t i})$ in succeeding cycles. Eventually one of the three conditions will hold for every column.

5. Analysis for decreasing u: The analysis here follows a pattern similar to that of the analysis for increasing u. We note that if we are interested in analysis for both increasing and decreasing u the operations of 4a. and 4b. can be carried out simultaneous with the operations of 5a. and 5b. respectively. The cells to be used in 4c., etc. (5c., etc.) are the ones containing positive (negative) entries in the table formed by doing 4b. and 5b. together.

5a. From the result of 3. we form a JxI table with entries

$$\begin{array}{lll} d_{j_o i} - d_{ji} & \text{if} & d_{j_o i} - d_{ji} < 0 \\ 0 \text{ (or a blank)} & \text{if} & 0 \leq d_{j_o i} - d_{ji} \end{array}$$

5b. With the results of 2. and 5a. we form a JxI table with entries

$$(c_{ji} - c_{j_0 i})/(d_{j_0 i} - d_{ji}) \quad \text{if} \quad d_{j_0 i} - d_{ji} < 0$$

$$-\infty \text{ (or a blank)} \quad \text{if} \quad 0 \leq d_{j_0 i} - d_{ji}$$

This is done by dividing each entry in the table of 2. by the corresponding entry in the table of 5a., adopting the convention that any number divided by zero is minus infinity.

5c. For each column in the table of 5b. we mark the maximum entry. This entry is u_{-1i} and we denote the cell in which it appears by $(j_{-1}i)$. We then enter this number in the lower right part of cell $(j_0 i)$ and the upper left part of cell $(j_{-1}i)$ in Table 1.

5d(i). We form a new table from the result of 5a. by subtracting from every entry in column i the entry in cell $(j_{-1}i)$. If the remainder is negative enter it in cell (ji) of the new table. Otherwise enter zero or leave the cell blank. In this way we form a table having as its entries

$$d_{j_{-1}i} - d_{ji} = (d_{j_0 i} - d_{ji}) - (d_{j_0 i} - d_{j_{-1}i}) \quad \text{if } d_{j_{-1}i} - d_{ji} < 0$$

$$0 \text{ (or a blank)} \quad \text{if } 0 \leq d_{j_{-1}i} - d_{ji}$$

5d(ii). Form a new table from the result of 2. by subtracting from each entry in column i the entry in cell $(j_{-1}i)$. We thus form a table with entries

$$c_{ji} - c_{j_{-1}i} = (c_{ji} - c_{j_0 i}) - (c_{j_{-1}i} - c_{j_0 i})$$

As will be seen in 5d(iii) below this operation need not be executed for those cells (ji) for which the corresponding entries in the result of 5d(i) are nonnegative.

5d(iii). From the results of 5d(i) and 5d(ii) we form a table containing the entries

$$\begin{array}{lll} (c_{ji} - c_{j_{-1}i})/(d_{j_{-1}i} - d_{ji}) & \text{if} & d_{j_{-1}i} - d_{ji} < 0 \\ -\infty \text{ (or a blank)} & \text{if} & 0 \leq d_{j_{-1}i} - d_{ji} \end{array}$$

This is accomplished by dividing each entry in the table of 5d(ii) by the corresponding entry in the table of 5d(i), adopting the convention that any number divided by zero is minus infinity.

5d(iv). We mark the maximum entry in each column of the table of 5d(iii). This number is u_{-2i} and the cell in which it appears is denoted by $(j_{-2}i)$. We then enter this number in the lower right part of cell $(j_{-1}i)$ and upper left part of cell $(j_{-2}i)$ in Table 1. Note that we now have specified in cell $(j_{-1}i)$ the range of u in which $c_i^* = c_{j_{-1}i}^*$ for $i = 1,\ldots,I$.

5d(v). We then begin a new cycle starting with the subtraction of the number in cell $(j_{-2}i)$ from every entry in column i for $i = 1,\ldots,I$ in the tables of 5d(i) and 5d(ii). The procedure is continued until at some cycle, -t, we find one of the following conditions holding for some column i.

5d(va). If we have a preassigned lower limit k_i on u we may find that $u_{-ti} \leq k_i$ indicating that the change in the subscript associated with c_i^* from j_{-t+1} to j_{-t} occurs outside (or on the boundary) of the interval within which u must lie. In this case we enter k_i rather than u_{-ti} in the lower right part of cell $(j_{-t+1}i)$ in Table 1.

5d(vb). It is possible at some cycle before condition (a) appears for us to encounter the condition

$$u_{-ti} = \frac{c_{j_{-t}i} - c_{j_{-t+1}i}}{d_{j_{-t+1}i} - d_{j_{-t}i}} > k_i \geq \underset{j=j_{-t}}{\text{Max}} \left(\frac{c_{ji} - c_{j_{-t+1}i}}{d_{j_{-t+1}i} - d_{ji}} \middle| \; d_{ji} > d_{j_{-t+1}i} \right)$$

The change in the subscript associated with c_i^* from j_{-t+1} to j_{-t} occurs at an admissible value of u, but the fact that the second largest entry in column i is exceeded by k_i implies that if u_{-t-1i} is computed it will be found to be outside (or on the border) of the interval within which u must lie. This follows from the statement in Section I A that

$$u_{-t-1i} \leq \frac{c_{j_{-t-1}i} - c_{j_{-t+1}i}}{d_{j_{-t+1}i} - d_{j_{-t-1}i}}$$

In this case we enter k_i rather than u_{-t-1i} in the lower right part of cell $(j_{-t}i)$ in Table 1.

5d(vc). Whether or not we have a lower limit on u a sufficiently large number of cycles will bring us to a cycle such that the subtraction of the number in cell $(j_{-t}i)$ will drive every entry nonnegative in column i of a table such as the one in 5d(i) above. This indicates that $d_{j_{-t}i} \geq d_{ji}$ for all j. In this case we enter $-\infty$ (or k_1 if such a lower limit is specified) in the lower right part of cell $(j_{-t}i)$ in Table 1 to indicate that $c_i^* = c_{j_{-t}i}^*$ for all admissible $u \leq u_{-ti}$.

When any of the conditions hold for a column no further analysis need be carried out on that column and we may simply leave out the corresponding columns in the tables for $(d_{j_{-t}i} - d_{ji})$ and $(c_{ji} - c_{j_{-t}i})$ in succeeding cycles. Eventually one of the three conditions will hold for every column.

In presenting the analysis procedure we have assumed implicitly that the subscripts j_t and j_{-t} are uniquely determined for each column. As noted previously this need not be true. However if at some cycle j_t or j_{-t} is not uniquely determined, we may use any of the eligible candidates to make the next transformation taking care to enter both u_{ti} and u_{t+1i} or u_{-ti} and u_{-t-1i} in the cells of Table 1 for every candidate.

Upon completing the analysis procedure above we are in a position to determine the optimal strategies for any admissible

value of u by inspection of Table 1. Given a value of u, say u^*, we mark the cell (ji) such that the two numbers in the cell bracket (in the sense of ordinary inequalities) u^* for every column i of Table 1. We may then construct the transpose of the optimal (for $u = u^*$) strategy matrix, $(S_o^*)'$, by placing ones in the cells of a JxI matrix which correspond to the marked cells in Table 1 and zeros elsewhere. If more than one cell of a column of Table 1 contains entries which bracket u^* in the sense of ordinary inequalities, there are alternative optimal strategies for $u = u^*$.

In general the expected loss of a strategy s_k is given by the equation

$$E\ (s_k) = tr\ (S_k C^*) = tr\ (S_k C) + u\ tr\ (S_k D)$$

and in particular the minimum expected loss when $u = u^*$ is

$$E\ (s_o^*) = tr\ (S_o^* C) + u^*\ tr\ (S_o^* D)$$

This minimum expected loss usually will be most easily computed if the matrix operations are executed first.

In some problems we will want to tabulate the strategies which are optimal for all admissible values of u. The following procedure permits us to do this in a reasonably compact form when we represent each strategy s_k by an I-tuple the i-th element of which is the subscript $j = 1, \ldots, J$ of the row in which the one occurs in the transpose of the strategy matrix S_k'.

We first rank the u_{ti}'s for all t and i. Say there are V + 1 such critical values which we may denote now by $u_1 < u_2 < \ldots < u_{V+1}$. If we have preassigned limits $u_1 = k_1$ and $u_{V+1} = k_2$. Otherwise we state formally that $u_1 = -\infty$ and $u_{V+1} = \infty$. We let the first entry in row v of a Vx(I+1) table be u_v for v = 1,...,V and we note u_{V+1} at the foot of the first column of this table. In the remaining I columns of row v we enter the I-tuple representing the optimal strategy for u such that $u_v \leq u \leq u_{v+1}$ for v = 1,...,V. This I-tuple is obtained by an inspection of Table 1. Some special provision such as expanding row v to two or more rows will be required to record alternative optimal strategies when they exist for all u such that $u_v \leq u \leq u_{v+1}$. Of course when u $= u_v$ the strategies represented in rows v-1 and v are both optimal.

While this procedure yields a list of all strategies which are optimal for some admissible value of u, it does not tell us what to do with the information in terms of selecting a strategy for use or restricting u more stringently. In this regard the ideas discussed in Chapter 3 remain valid in the one dimensional case although here the problems encountered may be somewhat more tractable by virtue of our dealing with only one parameter.

I C. Examples of Sensitivity Analysis: In this section we consider two examples of sensitivity analysis in the absence of simplifying assumptions of the type to be introduced later in this chapter.

I C 1. Example 1: We first consider an example which, although trivial in terms of analysis (the result is obvious from an examination of the criterion matrix in Chapter 1), is of some importance in finite statistical decision theory. We presume that the loss matrix can be written as

$$L^* = L + u E$$

where we require that $E = L$ and u be greater than -1. Following the procedure outlined in Chapter 1 we obtain the matrix from which the optimal strategy may be determined as

$$C^* = C + u C = (1 + u) C = h C$$

where $C = L \phi P$ and u greater then -1 implies that h is positive. The inspection procedure for the determination of the optimal strategy matrix for a given value of h entails finding the minimum element in each column of C^*, but for positive h we have

$$\underset{j}{\text{Min}}\ c^*_{ji} = h\ \underset{j}{\text{Min}}\ c_{ji}$$

Therefore the optimal strategy is completely insensitive to

variations in h. We may restate this result by saying that the multiplication of every element of the loss matrix by a positive constant will not alter the optimal strategy. Of course variations in h influence the minimum expected loss because

$$E(s_o) = tr(S_o C^*) = h\, tr(S_o C)$$

I C 2. Example 2: This example is presented to illustrate the analysis procedure of section I B and the handling of probability distribution information in the form of a marginal distribution over indications and a posterior distribution over states of nature.

We consider a simple inventory system for a commodity subject to such deterioration that any unit held for longer than one week must be discarded. We presume that requirements for the commodity occur only during the five-day period Monday through Friday and that the demand process can be represented as a Markov chain in equilibrium. More precisely, $\theta_n = n-1$ for $n=1,\ldots,6$ are the requirements to be experienced "next" week and are related to the requirements experienced "this" week, $x_i = i-1$ for $i=1,\ldots,6$ by a transition or posterior distribution matrix, Q, giving Prob $(\theta_n = n-1 \mid x_i = i-1)$.

$$\begin{array}{cccccccccc}
 & & & & x_i = i-1 & & & & & \\
 & 0 & 1 & 2 & 3 & 4 & 5 & & & \\
 & .50 & .30 & .10 & 0 & 0 & 0 & & 0 & \\
 & .30 & .40 & .25 & .10 & 0 & 0 & & 1 & \\
Q = & .20 & .20 & .30 & .25 & .10 & 0 & & 2 & \theta_n = n-1 \\
 & 0 & .10 & .25 & .30 & .20 & .20 & & 3 & \\
 & 0 & 0 & .10 & .25 & .40 & .30 & & 4 & \\
 & 0 & 0 & 0 & .10 & .30 & .50 & & 5 &
\end{array}$$

The prior distribution, Prob ($x_i = i-1$), of the indications is

$$(.14,\ .18,\ .18,\ .18,\ .18,\ .14)$$

and because the process is in equilibrium it is also the distribution, Prob ($\theta_n = n-1$), if we have no information about the requirements experienced in preceding weeks. We can arrange the prior distribution in a diagonal matrix, $\mathcal{P}$, and compute the joint distribution Prob ($\theta_n = n-1$, $x_i = i-1$).

$$Q\mathcal{P} = \begin{pmatrix}
.070 & .054 & .018 & 0 & 0 & 0 \\
.042 & .072 & .045 & .018 & 0 & 0 \\
.028 & .036 & .054 & .045 & .018 & 0 \\
0 & .018 & .045 & .054 & .036 & .028 \\
0 & 0 & .018 & .045 & .072 & .042 \\
0 & 0 & 0 & .018 & .054 & .070
\end{pmatrix}$$

On Friday evening the requirements for "this" week are noted and an order is placed for $a_j = j-1$ ($j = 1,\ldots,6$) units for Monday morning delivery. We never order more than five units because any excess over maximum demand would surely be lost to deterioration. We suppose that the objectives

governing the decision are

O_1: Minimize the number of units not used during the week. We can express the measure of efficiency relative to O_1 as

$$z_1 = \begin{array}{ll} a_j - \theta_n = j - n & \text{if } \theta_n \leq a_j \\ 0 & \text{if } \theta_n > a_j \end{array}$$

We take the scale of measurement of z_1 as our standard loss scale, i.e. $u_1(z_1) = z_1$.

O_2: Minimize the number of shortages (requirements that cannot be fulfilled out of stock) during the week. We can express the measure of efficiency relative to O_2 as

$$z_2 = \begin{array}{ll} 0 & \text{if } \theta_n \leq a_j \\ \theta_n - a_j = n - j & \text{if } \theta_n > a_j \end{array}$$

We assume that a procedure such as the one discussed in Chapter 2 (section III A 2c) yields an estimate of $u_2(z_2)$ in the form of two bracketing polynomials which is equivalent to

$$u_2\,(z_2) = z_2 + u_{22}\,(z_2^2 - .1z_2^3) \quad \text{for } 0 \leq z_2 \leq 5$$

where $1 \leq u_{22} \leq 2$,

and that we use an average value of u_{22} to express the value function which transforms z_2 into the standard

units as

$$u_2\,(z_2) = z_2 + 1.5\,(z_2^2 - .1\,z_2^3) \quad \text{for } 0 \leq z_2 \leq 5$$

The construction of the loss matrix now consists of entering the appropriate values of z_1 or $u_2(z_2)$ in each cell.

$$\theta_n = n-1$$

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & & \\ \left(\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}\right. & \begin{array}{c} 2.35 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{array}{c} 6.80 \\ 2.35 \\ 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{array}{c} 12.95 \\ 6.80 \\ 2.35 \\ 0 \\ 1 \\ 2 \end{array} & \begin{array}{c} 18.40 \\ 12.95 \\ 6.80 \\ 2.35 \\ 0 \\ 1 \end{array} & \left.\begin{array}{c} 23.75 \\ 18.40 \\ 12.95 \\ 6.80 \\ 2.35 \\ 0 \end{array}\right) & \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & a_j = j-1 \end{array}$$

To determine the optimal strategy we first compute $C = LQ'$ and then find the minimum element in each column of C (indicated by underscoring below).

$$x_1 = i-1$$

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & & \\ \left(\begin{array}{c} .289 \\ \underline{.136} \\ .182 \\ .322 \\ .462 \\ .602 \end{array}\right. & \begin{array}{c} .647 \\ .261 \\ \underline{.223} \\ .342 \\ .522 \\ .702 \end{array} & \begin{array}{c} 1.387 \\ .634 \\ .309 \\ \underline{.240} \\ .360 \\ .540 \end{array} & \begin{array}{c} 2.303 \\ 1.387 \\ .634 \\ .309 \\ \underline{.240} \\ .360 \end{array} & \begin{array}{c} 3.196 \\ 2.213 \\ 1.274 \\ .544 \\ .199 \\ \underline{.198} \end{array} & \left.\begin{array}{c} 2.798 \\ 2.022 \\ 1.258 \\ .575 \\ .193 \\ \underline{.098} \end{array}\right) & \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & a_j = j-1 \end{array}$$

The optimal strategy, s_o, can be represented by the sextuple

(1, 2, 3, 4, 5, 5)

It incorporates a safety factor against shortages by ordering

for a "next" week requirement one unit greater than that observed "this" week. The minimum expected loss is

$$E\ (s_o) = \text{tr}\ (S_o C) = 1.135$$

Because this strategy is based on an estimate $u_{22} = 1.5$, questions concerning its validity may arise. We might ask how far u_{22} can vary (within the limits of the bracketing conditions) from 1.5 before this strategy ceases to be optimal. To answer such a question we may let $u = u_{22} - 1.5$ represent variations of u_{22} from its mid-range value in which case the bracketing restrictions on u_{22} become

$$-.5 \leq u \leq .5$$

We express the loss matrix as the following linear function of u.

$$L^* = L + u\ E$$

where E is

$$\begin{array}{c} \\ \\ E = \end{array} \begin{array}{cccccc|cc} \multicolumn{6}{c}{\theta_n = n-1} & & \\ 0 & 1 & 2 & 3 & 4 & 5 & & \end{array}$$

$$E = \left(\begin{array}{cccccc} 0 & .9 & 3.2 & 6.3 & 9.6 & 12.5 \\ 0 & 0 & .9 & 3.2 & 6.3 & 9.6 \\ 0 & 0 & 0 & .9 & 3.2 & 6.3 \\ 0 & 0 & 0 & 0 & .9 & 3.2 \\ 0 & 0 & 0 & 0 & 0 & .9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \quad a_j = j-1$$

because variations in u_{22} influence loss matrix entries via the multiplier $(z_2^2 - .1z_2^3)$.

Before beginning the sensitivity analysis we express the matrix from which the optimal strategy is derived as a linear function of u.

$$C^* = C + u\,D$$

$$x_i = i-1$$

$$D = Maf = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \\ \end{array}$$

$$\begin{array}{c} \\ D = Maf = \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \left(\begin{array}{cccccc} .127 & .293 & .670 & 1.157 & 1.651 & 1.455 \\ \underline{.025} & .090 & .306 & .670 & 1.103 & 1.026 \\ 0 & \underline{.016} & .098 & .306 & .603 & .601 \\ 0 & 0 & \underline{.016} & .098 & .238 & .262 \\ 0 & 0 & 0 & \underline{.016} & .049 & .063 \\ 0 & 0 & 0 & 0 & \underline{0} & \underline{0} \end{array}\right) \end{array} \begin{array}{l} \\ 0 \\ 1 \\ 2 \quad a_j = j-1 \\ 3 \\ 4 \\ 5 \end{array}$$

Ignoring the underscored elements in D for the moment we begin the sensitivity analysis procedure with step 2 deferring the presentation of Table 1 until later.

In step 2 we subtract the underscored elements in C above from all elements in their respective columns to obtain the following table.

.153	.425	1.147	2.063	2.998	2.700
0	.039	.444	1.147	2.015	1.924
.046	0	.069	.444	1.076	1.160
.186	.120	0	.069	.356	.477
.326	.300	.120	0	.001	.095
.466	.480	.300	.120	0	0

In step 3 we subtract the underscored elements in D above from all elements in their respective columns and reverse the signs of the remainders to obtain the following table.

-.102	-.277	-.653	-1.141	-1.651	-1.455
0	-.074	-.290	- .653	-1.103	-1.026
.025	0	-.082	- .290	- .603	- .601
.025	.016	0	- .082	- .238	- .262
.025	.016	.016	0	- .049	- .063
.025	.016	.016	.016	0	0

Because we will analyze for both increasing and decreasing u, we combine steps 4b and 5b incorporating 4a and 5a later. We divide each entry in the table of step 2 by the corresponding entry in the table of step 3 to obtain the following table.

-1.500	-1.532	-1.755	-1.808	-1.816	-1.856
±∞	-.524	-1.531	-1.754	-1.826	-1.875
1.893	±∞	-.841	-1.531	-1.724	-1.931
7.389	7.389	±∞	-.841	-1.500	-1.821
12.944	18.500	7.389	±∞	-.185	-1.500
18.500	29.611	18.500	7.389	±∞	±∞

Considering the analysis for increasing u first, we note that step 4a would have led to the assignment of ∞ to every cell containing a nonpositive entry in the table above. Upon examination of the positive entries in the table we find that all of them exceed .5, the upper limit on u. Therefore by condition 4d(va) the optimal strategy will not change for positive admissible values of u, and we make the entry .5 in the upper left part of the appropriate cells in Table 1.

Step 5a in the analysis for decreasing u would have led to the assignment of -∞ to every cell containing a nonnegative entry in the table above. Upon examination of the negative

entries, we find that except for the fifth column the maximum entry in each column (indicated by underscoring) is less than -.5, the lower limit on u. Therefore by condition 5d(va) we enter -.5 in the lower right part of the appropriate cells in Table 1. The second largest value in the fifth column is -1.5. Therefore by condition 5d(vb) no other change in the optimal strategy will occur for negative admissible values of u; we enter -.185 in the lower right part of cell (4,4) and upper left part of cell (4, 3) and enter -.5 in the lower right part of cell (4, 3) in Table 1.

Table 1

$x_i = i-1$

0	1	2	3	4	5	
						0
.5 -.5						1
	.5 -.5					2
		.5 -.5				3
			.5 -.5	-.185 -.5		4
				.5 -.185	.5 -.5	5

$a_j = (j-1)$

We also may summarize the strategies which are optimal for all admissible values of u as follows.

Strategy Designation	Critical Values u	u_{22}	Strategy Description
s_1	-.5	1.0	(1,2,3,4,4,5)
s_2	-.185	1.315	(1,2,3,4,5,5)
	.5	2.0	

That is for $-.5 \leq u \leq -.185$ s_1 is optimal and for $-.185 \leq u \leq .5$ s_2 is optimal.

The range of expected losses which may be encountered using either of the two strategies depends on u.

For u = -.5				For u = .5
1.075	$\leq$	$E(s_1)$	$\leq$	1.197
1.098	$\leq$	$E(s_2)$	$\leq$	1.172

Given this information on the sensitivity of the optimal strategy to variations in the polynomial value function, $u_2(z_2)$, within the limits of the bracketing relations, we still have to decide whether to simply select a strategy for use now or to attempt to restrict $u_2(z_2)$ more stringently.

If we decide to forego additional analysis and to use the strategy with the largest optimality subset, we will select s_2 since it is optimal over a longer interval than is s_1. If on the other hand we decide to further restrict

$u_2(z_2)$, the results of the analysis make it seem reasonable to first try to obtain a new $u_2^*(z_2)$ polynomial. For example if we can alter $u_2^*(z_2)$ from $z_2 + z_2^2 - .1z_2^3$ to $z_2 + 1.5z_2^2 - .15z_2^3$ we will have restricted $u_2(z_2)$ enough so that only a_2 could be optimal. Of course a new $u_2^*(z_2)$ might not materialize in such a convenient form, and we might have to again go through the process of reducing the bracketing relations to linear restrictions on the coefficients of $u_2(z_2)$ followed by another analysis conforming to these new restrictions.

I C 3. Example 3: When the parameter under consideration influences only the losses experienced under a particular state of nature, say θ_a, only the a-th column of E contains nonzero entries. Thus we can write the entries of D as

$$d_{ji} = \sum_{n=1}^{n=N} e_{jn} q_n p_{ni} = e_{ja} q_a p_{ai} \quad \text{for } j = 1,\ldots,J \; ; \; i = 1,\ldots,I$$

and can express the general statements of section A as

$$u_{t+1\,i} = (q_a p_{ai})^{-1} \frac{e_{j_{t+1}i} - e_{j_t i}}{e_{j_t a} - e_{j_{t+1}a}} = (q_a p_{ai})^{-1} \underset{j}{\text{Min}} \left(\frac{e_{ji} - e_{j_t i}}{e_{j_t a} - e_{ja}} \;\middle|\; e_{ja} < e_{j_t a} \right)$$

$$u_{\underline{t}-1\,i} = (q_a p_{ai})^{-1} \frac{e_{\underline{j}_{t-1}i} - e_{\underline{j}_t i}}{e_{\underline{j}_t a} - e_{\underline{j}_{t-1}a}} = (q_a p_{ai})^{-1} \underset{j}{\text{Max}} \left(\frac{e_{ji} - e_{\underline{j}_t i}}{e_{\underline{j}_t a} \; e_{ja}} \;\middle|\; e_{\underline{j}_t a} < e_{ja} \right)$$

for $t = 0, 1, \ldots.$

This assumption in itself does not bring any simplifications of the analysis procedure within reach, but when it is supplemented by the specification $e_{ja} = 1$ $(j = 1, \ldots, J)$ we have a case that is of some importance in finite statistical decision theory, the addition of an arbitrary constant to the elements of the loss matrix which correspond to a particular state of nature. If we let $t = 0$ in the two expressions above we will find that the criteria for u_{1i} and u_{-1i} cannot be computed because there is no j such that e_{ja} is less than e_{j_0a} or e_{j_0a} is less than e_{ja}. In section 4 we suggested that this situation could be formally represented by $u_{1i} = \infty$ and $u_{-1i} = -\infty$. In other words the optimal strategy is totally insensitive to variations in a parameter u such that

$$L^* = L + u E$$

where one column of E contains ones while all other elements are zero.

Clearly we can repeat this analysis for each state of nature and draw the conclusion that the optimal strategy will not be altered if we replace l_{jn} by $l_{jn} + g(\theta_n)$ for $j = 1, \ldots, J$ and $n = 1, \ldots, N$ where $g(\theta_n)$ is a function only of the state of nature. In particular we can select $g(\theta_n)$ such that

$$-g(\theta_n) = \operatorname*{Min}_{j} l_{jn} \qquad \text{for } n = 1, \ldots, N$$

and can work with what we may call, following Chernoff and Moses (1959), "regrets," i.e. the losses incurred by taking actions, a_j, in the presence of states of nature, θ_n, over and above the minimum losses attainable under θ_n. With this formulation the y_{kn} of Chapter 1 (section II) are referred to as "risks" and $E(s_o)$ may be called the "minimum expected risk."

Furthermore we can combine the result of this example with that of example 1 to produce the statement that the optimal strategy will not be altered if we replace l_{jn} by $g(\theta_n) + h\, l_{jn}$ $(h > 0)$ for $j = 1,\ldots,J$ and $n = 1,\ldots,N$. In particular when we set $g(\theta_n) = g$ where g is an arbitrary constant, we have the statement that it is sufficient (although not always necessary) to specify the loss matrix unique up to a positive linear transformation in order to determine an optimal strategy.

II. Analysis When D Contains Non-zero Elements in Only One Row: In this section we consider the sensitivity of the optimal strategy of the finite statistical decision problem to variations in a parameter which affects only one row of D.

Such a parameter could be one that influences only the losses associated with a particular terminal action. Then only one row of E (and consequently only one row of $D = E/P$) contains non-zero elements. Two special cases may sometimes be of interest. One is the case when every element of the non-zero row of E is unity which represents analysis of the

sensitivity of the optimal strategy to the addition of an arbitrary constant to every loss corresponding to a particular terminal action. The other is the case when the non-zero row of D is a unit vector which represents analysis for variations in one element of the loss matrix.

II A. Analysis of Changes in the Optimal Strategy: The structure of the C^* matrix used to specify the optimal strategy is such that the analysis procedure can be simplified when D contains non-zero entries in only one row. In the general discussion we will assume that the non-zero row of D is arbitrary.

II A 1. Preliminary discussion: It will be convenient for purposes of discussion to assume that the rows of C and D have been labelled so that only the first row of D contains non-zero elements. With this labelling we can write the elements of $C^* = C + uD$ as

$$c^*_{ji} = c_{ji} + ud_{ji} = \begin{array}{ll} c_{1i} + ud_{1i} & \text{for } j = 1 \\ c_{ji} & \text{for } j = 2,\ldots,J \end{array}$$

Therefore only the elements in the first row of C^* vary with u, and the question of the subscript j associated with

$$c^*_i = \min_j c^*_{ji} \qquad \text{for } i = 1,\ldots,I$$

simplifies to the question of whether or not

$$c^*_{1i} \leq c_{j_1 i} \quad \text{where } c_{j_1 i} = \min_{1<j} c_{ji}$$

for various values of u. There are at most two subscripts associated with c_i^* for any u, i and j_i. If a change from one subscript to the other takes place at all, it occurs at either $u = u_{ii}$ or $u = u_{-ii}$. Whether a change does occur is found by examining the expressions for u_{ii} and u_{-ii} of section I A which here simplify to

$$\frac{c_{j_i i} - c_{ii}}{d_{ii}} = u_i = \begin{array}{ll} u_{ii} & \text{if } u_i \text{ is non-negative} \\ u_{-ii} & \text{if } u_i \text{ is non-positive} \end{array}$$

A tedious but straightforward examination of this expression for various values of d_{ii}, c_{ii}, and $c_{j_i i}$ produces the information needed to specify the subscript associated with c_i^* for any value of u. This information is summarised in the following table.

u_{-ii} is non-positive				u_{ii} is non-negative		
d_{ii} is				d_{ii} is		
<0	$=0$	>0		<0	$=0$	>0
u_i^{+}	☒	☒	$c_{j_i i} > c_{ii}$	☒	∞^{++}	u_i^{++}
0^{+}	☒	0^{++}	$c_{j_i i} = c_{ii}$	0^{+}	☒	0^{++}
☒	$-\infty^{++}$	u_i^{++}	$c_{j_i i} < c_{ii}$	u_i^{+}	☒	☒

If d_{ii}, c_{ii}, and $c_{j_i i}$ are such that u_i falls in a cell of the table marked by the superscript

(+) as u increases through u_i^{+} the subscript associated with c_i^* changes from j_i to i. That is

$$c_i^* = \begin{cases} c_{j_i i} & \text{for} \quad u \leq u_i^{+} \\ c_{ii}^* & \text{for} \quad u \geq u_i^{+} \end{cases}$$

(++) as u increases through u_i^{++} the subscript associated with c_i^* changes from i to j_i. That is

$$c_i^* = \begin{cases} c_{ii}^* & \text{for} \quad u \leq u_i^{++} \\ c_{j_i i} & \text{for} \quad u \geq u_i^{++} \end{cases}$$

When u_i is non-zero it will fall in one and only one cell of the table. In particular if $d_{ii} = 0$ and $c_{j_i i} - c_{ii} \neq 0$, $u_i = \pm \infty$ according to the convention that a positive (negative) number divided by zero is plus (minus) infinity. The infinity symbols signify that there will be no change in the subscript associated with c_i^* as u varies arbitrarily. If u_i is zero it will fall in two cells of the table, one on the left and one on the right, representing the fact that $u_{ii} = u_{-ii} = 0$. If both numerator and denominator of the expression for u_i are zero, $u_i = 0/0$ is indeterminate and there is no cell in the table open to accomodate it. In the context of sensitivity analysis this means that there will be no change in the subscript associated with c_i^* as u varies arbitrarily because $c_i^* = c_{ii} = c_{j_i i}$ for all u.

II A 2. Analysis procedure: We may use the results of the preceding discussion to devise a series of manipulations on the C matrix that will convert it to a form such that we can immediately specify an optimal strategy for any value of u. Given the matrices C and D we may proceed as follows.

a. In each column of C we mark (e.g. underscore) $c_{j_i i} = \min\limits_{i<j} c_{ji}$

b. For each column of C we compute $u_i = (c_{j_i i} - c_{ii})/d_{ii}$ and enter this number at the foot of column i marked by a superscript + or ++ according to the position of this u_i in the table of section 1. If $u_i = 0/0$ we need make no entry.

c. We rearrange the columns of C (i.e. relabel the indications) so that the row at the foot of C can be divided into four groups as is indicated below.

Group No.	i			ii			iii			iv		
Col. No.	1	...	I_1	I_1+1	...	I_2	I_2+1	...	I_3	I_3+1	...	I
u_i	$-\infty$	...	$-\infty$	u_{I_1+1}	...	u_{I_2}	∞	...	∞	0/0	...	0/0

The columns within group ii are ordered so that, $-\infty < u_{I_1+1} \leq u_{I_1+2} \leq \dots \leq u_{I_2} < \infty$. The ordering of the columns within groups i, iii, and iv is arbitrary.

If group i is empty we set $I_1 = 0$ and define $u_0 = -\infty$ to aid in formally specifying an optimal strategy when u is less than u_1. If group iii is empty we set $I_3 = I_2$ and define $u_{I_2+1} = \infty$ for purposes of specifying an optimal strategy

when u is greater than u_{I_2}. If both groups iii and iv are empty we set $I_2 = I$. We do not consider the trivial case when group ii is empty, i.e. $D = 0$ and the optimal strategy is totally insensitive to variations in u.

d. For the columns of C with u_i marked by the superscript +(++) mark c_{1i} by the superscript +(++) and $c_{j_1 i}$ by the superscript ++(+).

After the completion of this procedure we may specify the optimal strategy matrix for any given value of u as follows. We find r and t ($I_1 = r$, $1 = t$, $r+t = I_2 + 1$) such that

$$u_r < u < u_{r+t}$$

Our use of strict inequalities here implies that

$$u = u_{r+1} = \dots = u_{r+t-1}$$

We form the transpose of the optimal strategy matrix by placing ones in the following cells of a JxI matrix.

a. In cells corresponding to the superscripts + in C for columns $1,\dots,r$ (not applicable if $r = 0$)

b. In cells corresponding to the superscripts ++ in C for columns $r+t,\dots,I_3$ (not applicable if $r+t = I_3+1 = I_2+1$).

c. In cells corresponding to either the superscripts + or ++ in C for columns $r+1,\dots,r+t-1$ (not applicable if $t = 1$).

d. In either cell (i,i) or (j_i,i) in columns for which u_i is indeterminate (columns $I_3+1,\ldots,I$).

The transpose of the optimal strategy matrix is completed by entering zeros in the remaining cells of the $J \times I$ matrix. We note that if either c. or d. above is applicable there are alternative optimal strategies for the given value of u.

II B. The Minimum Expected Loss: The expected loss of any strategy, s_k, is the following linear function of u.

$$E\ (s_k) = \text{tr}\ (S_k C) + u\ \text{tr}\ (S_k D)$$

In particular if s_{oi} is one of the optimal strategies when $u_i \leq u \leq u_{i+1}$ $(I_1 \leq i \leq I_2)$ the minimum expected loss is the following piece-wise linear function of u.

$$E\ (s_o) = E\ (s_{oi}) \quad \text{for } u_i \leq u \leq u_{i+1}\ .\ I_1 \leq i \leq I_2$$

In section I we found it necessary to compute $\text{tr}(S_{oi}C)$ and $\text{tr}(S_{oi}D)$ by matrix operations for each i in order to determine this function over some interval. The nature of the problem in this section is such that the use of another method can save considerable effort if we wish to express $E(s_o)$ over an interval in which several strategies are optimal.

II B 1. Preliminary discussion: To develop this method we first present some relations among $\text{tr}(S_{ov}C)$, $\text{tr}(S_{ov+1}C)$,

$tr(S_{ov}D)$, and $tr(S_{ov+1}D)$ which hold after completing the analysis procedure of section A.

The ordering of the columns of D is the same as the ordering of the columns of C. We may attach the same superscript (+ or ++) to d_{iv} as is attached to u_v and c_{iv}. Then because of the manner in which S_{ov} is constructed, we can express $tr(S_{ov}D)$ as

$$tr\,(S_{ov}D) = \sum_{i=I_1+1}^{i=v} d_{ii}^{+} + \sum_{i=v+1}^{i=I_2} d_{ii}^{++} \quad \text{for } I_1 \leq v \leq I_2$$

That is we sum the elements of D marked by the superscript + for columns I_1+1 through v and those marked by ++ for columns v+1 through I_2. Because $d_{ji} = 0$ for $i < j$ and $d_{ii} = 0$ for $i \leq I_1$ and $I_2 < i$, these elements of D do not appear. We drop the first summation above to obtain $tr(S_{oI_1}D)$ and the second summation to obtain $tr(S_{oI_2}D)$.

When u_v is less than u_{v+1} the equation above follows directly from the structure of S_{ov}. When $u_v = u_{v+1}$ we define s_{ov} to be that strategy among strategies optimal for $u = u_v = u_{v+1}$ for which the equation holds.

The following relation can be distilled from the results of examination of $tr(S_{ov+1}D) - tr(S_{ov}D)$ for various combinations of superscripts on $d_{iI_1+1}, \ldots, d_{iI_2}$.

$$tr\,(S_{ov+1}D) = tr\,(S_{ov}D) + g_{v+1} \quad \text{for } I_1 \leq v \leq I_2-1$$

where

$$s_{v+1} = \begin{cases} d_{1v+1} & \text{if } u_{v+1} \text{ has the superscript } + \\ -d_{1v+1} & \text{if } u_{v+1} \text{ has the superscript } ++ \end{cases}$$

Similarly because of the manner in which S_{ov} is constructed we can express $tr(S_{ov}C)$ as

$$tr\,(S_{ov}C) = \sum_{i=1}^{i=v} (c_{1i}^{+} + c_{j_1 i}^{+}) + \sum_{i=v+1}^{i=I_3} (c_{1i}^{++} + c_{j_1 i}^{++}) + \sum_{i=I_3+1}^{i=I} c_{1i}$$

$$\text{for } I_1 \leq v \leq I_2$$

That is we sum the elements of C marked by the superscript + for columns 1 through v, those marked by ++ for columns v+1 through I_3, and the $c_{1i} = c_{j_1 i}$ for these columns having u_i indeterminate. When $I_3 = I$ we drop the third summation above. If $I_1 = 0$ we drop the first summation above to obtain $tr(S_{oo}C)$, and if $I_2 = I_3$ we drop the second summation above to obtain $tr(S_{oI_2}C)$.

As was noted in connection with the computation of $tr(S_{ov}D)$, when $u_v = u_{v+1}$ we have alternative optimal strategies to choose among for our computation of $E(s_o)$. We adopt the convention of using as s_{ov} that strategy for which the given equations for $tr(S_{ov}C)$ and $tr(S_{ov}D)$ hold.

Examination of $tr(S_{ov+1}C) - tr(S_{ov}C)$ for various combinations of superscripts on $d_{1I_1+1}, \ldots, d_{1I_2}$ yields results that

can be condensed into the following relation.

$$\mathrm{tr}\,(S_{ov+1}C) = \mathrm{tr}\,(S_{ov}C) + h_{v+1} \quad \text{for } I_1 \leq v \leq I_2 - 1$$

where

$$h_{v+1} = \begin{matrix} -\,(c_{j_{v+1}v+1} - c_{iv+1}) & \text{if } u_{v+1} \text{ has the superscript } + \\ (c_{j_{v+1}v+1} - c_{iv+1}) & \text{if } u_{v+1} \text{ has the superscript } ++ \end{matrix}$$

We further note that because $u_i = (c_{j_i i} - c_{ii})/d_{ii}$ we can relate g_i and h_i as follows.

$$h_i = -u_i g_i \quad \text{for } I_1 + 1 \leq i \leq I_2$$

II B 2. Computational procedure: The relations above suggest the following procedure when it is necessary to compute several $\mathrm{tr}(S_{oi}C)$ and $\mathrm{tr}(S_{oi}D)$ in order to represent $E(s_o)$ as a piece-wise linear function of u over some range of u.

We first use matrix operations to compute $\mathrm{tr}(S_{ov}C)$ and $\mathrm{tr}(S_{ov}D)$ for one strategy, s_{ov}, which is optimal somewhere in the interval with which we are concerned and then we construct a 5xI table at the foot of our C matrix the rows of which are determined as follows.

(a.) Row 1 contains the u_i's and is already determined by the analysis procedure of section A.

(b.) Row 2 contains the g_i's which are obtained using the D matrix and the specifications in section 1 above.

(c.) Row 3 contains the h_i's which are obtained by multiplying the entry in row 1 by the entry in row 2 and entering the result in row 3 with its sign reversed.

(d.) Row 4 contains the $tr(S_{oi}C)$'s which are obtained by first entering $tr(S_{ov}C)$ under u_v. Then $tr(S_{ov+1}C)$ is obtained as the sum of the entries in the cells on the left of and above the cell in row 4 under u_{v+1}. We may repeat this process for columns to the right of u_v as far as is desired. Then $tr(S_{0v-1}C)$ is obtained as the entry to the right of its cell minus the entry immediately above its cell. We may repeat this process for columns to the left of u_v as far as is desired.

(e.) Row 5 contains the $tr(S_{oi}D)$'s. We first enter $tr(S_{ov}D)$ under u_v. Then $tr(S_{ov+1}D)$ is computed by adding the entry in row 2 under u_{v+1} to the entry in row 5 under u_v. We may repeat this process for columns to the right of u_v as far as is desired. Then $tr(S_{0v-1}D)$ is computed by subtracting the entry in row 2 under u_{v-1} from the entry in row 5 under u_v. We may repeat this process for columns to the left of u_v as far as is desired.

We note that these computations need not be carried out in columns to the left of a column for which u_i is less than k_1 (or $u_i = -\infty$) or to the right of a column for which u_i is greater than k_2 (or $u_i = \infty$) where k_1 and k_2 are preassigned lower and upper limits on u.

II C. Example of Sensitivity Analysis (Example 4): As an illustration of two aspects of sensitivity analysis for variations in a single parameter we consider a problem suggested by Birnbaum and Maxwell (1960). The decision problem is to classify an arriving patient into one of the four broad psychiatric diagnostic categories below.

State of nature: the patient belongs in category	Prob (θ_n) q_n
θ_1: Neurotic	.461
θ_2: Schizophrenic	.237
θ_3: Affective disorder	.157
θ_4: Abnormal personality	.145
	1.000

We take as our prior distribution here the prior probabilities estimated by Birnbaum and Maxwell from a sample of patients entering two London hospitals.

The experimental procedure consists of determining the presence or absence of each of nine characteristics.

Theoretically there are $2^9 = 512$ diagnostic indications. In a sample of 772 patients reported by Birnbaum and Maxwell, 178 indications actually occurred while for purposes of exposition we will assume that only the ten following (a check indicates the presence of the characteristic) are possible.

Characteristics	Diagnostic Indications									
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
Hysterical symptoms, life long or episodic	✓					✓				
Anxiety	✓	✓					✓			
Schizophrenic type thought disorder				✓						
Depression			✓	✓	✓					✓
Compulsive acts and/or obsessional thoughts	✓	✓								✓
Lack of confidence when in society				✓	✓	✓		✓		
Definite mood variations before present illness			✓						✓	
Impulsiveness and/or aggressiveness									✓	
Hypochondriacal attitude toward illness			✓				✓	✓		

We will take the following matrix as our conditional probability distribution P. To emphasize that the numbers in this matrix do not represent real data, we note the data of Birnbaum and Maxwell would lead to the estimates (.0381, .0541, .0287, .0183) of the probabilities in the column under x_5.

Diagnostic Indications

$$P = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{cccccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} \end{array}$$

$$P = \left(\begin{array}{cccccccccc} .20 & .16 & .06 & .10 & .12 & .13 & .05 & .06 & .04 & .08 \\ .06 & .05 & .15 & .25 & .10 & .08 & .07 & .04 & .10 & .10 \\ .05 & .03 & .09 & .08 & .08 & .10 & .30 & .20 & .05 & .02 \\ .10 & .08 & .12 & .08 & .06 & .03 & .03 & .00 & .22 & .28 \end{array}\right)\begin{array}{ll} \theta_1 & \\ \theta_2 & \text{States} \\ \theta_3 & \text{of} \\ \theta_4 & \text{Nature} \end{array}$$

After arranging the prior distribution in a diagonal matrix, ϕ we may compute the joint distribution matrix.

$$\phi P = \begin{pmatrix} .092 & .074 & .039 & .046 & .055 & .060 & .023 & .028 & .018 & .037 \\ .014 & .012 & .036 & .059 & .024 & .019 & .017 & .009 & .024 & .024 \\ .008 & .005 & .014 & .013 & .013 & .016 & .047 & .031 & .008 & .003 \\ .015 & .012 & .017 & .012 & .009 & .004 & .004 & .000 & .032 & .041 \end{pmatrix}$$

We consider five actions, to classify the patient as being in one of the four categories assigning him to a clinic appropriate for that category, or to refuse to classify the patient pending further testing. We suppose that we accept the following as our loss matrix.

$$\begin{array}{cccccl} & \multicolumn{4}{c}{\text{States of Nature}} & \text{Actions} \\ & \theta_1 & \theta_2 & \theta_3 & \theta_4 & \\ & 0 & 1 & 1 & 1 & a_1\text{: classify as neurotic} \\ & 1 & 0 & 1 & 1 & a_2\text{: classify as schizophrenic} \\ L = & 1 & 1 & 0 & 1 & a_3\text{: classify as affective disorder} \\ & 1 & 1 & 1 & 0 & a_4\text{: classify as abnormal personality} \\ & u & u & u & u & a_5\text{: refuse to classify} \end{array}$$

where u is the expected cost of the procedure to be used if we refuse to classify the patient now which we express as a fraction of the cost of misclassifying a patient.

It is possible that rather than being a single number, u is a function of the indications, $u(x_i)$. For example if we

refuse to classify the patient now, the additional tests performed after observing x_1 may be different (and either more or less costly) than the tests after observing x_{10}. One reason for a sensitivity analysis is that if we find all $u(x_i)$ falling in one of the intervals within which the optimal strategy is insensitive to variations in u we can avoid formulating the much more complex model that is required if the relations between the indications and future test procedures are to be explicitly considered.

We recall from examples 1 and 3 that to specify an optimal strategy it is sufficient to measure the loss matrix elements unique up to a positive linear transformation. Thus we could replace zero and one in the loss matrix by v and w v respectively without altering the optimal strategy. Of course such a replacement would alter the origin and unit of measurement of the scale on which we measure u. We employ the numbers zero and one here as a matter of convenience.

We first observe that if $u = 1$ (or greater) a_5 will never be selected in an optimal strategy because a_5 dominates each other action in the sense $l_{jn} \leq l_{5n}$ for $j,n = 1,\ldots,4$. Let us set $u = 1$ and examine the optimal strategy derived from $C = L\phi P$.

refuse to classify the patient now, the additional tests performed after observing x_1 may be different (and either more or less costly) than the tests after observing x_{10}. One reason for a sensitivity analysis is that if we find all $u(x_i)$ falling in one of the intervals within which the optimal strategy is insensitive to variations in u we can avoid formulating the much more complex model that is required if the relations between the indications and future test procedures are to be explicitly considered.

We recall from examples 1 and 3 that to specify an optimal strategy it is sufficient to measure the loss matrix elements unique up to a positive linear transformation. Thus we could replace zero and one in the loss matrix by v and w   v respectively without altering the optimal strategy. Of course such a replacement would alter the origin and unit of measurement of the scale on which we measure u. We employ the numbers zero and one here as a matter of convenience.

We first observe that if u = 1 (or greater) a_5 will never be selected in an optimal strategy because a_5 dominates each other action in the sense $l_{jn} \leq l_{5n}$ for $j, n = 1, \ldots, 4$. Let us set u = 1 and examine the optimal strategy derived from $C = L/P$.

Indications

$$\begin{array}{cc|cccccccccc|c}
 & & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & \text{Actions} \\
 & & \underline{.037} & \underline{.029} & .067 & .084 & \underline{.046} & \underline{.039} & .068 & .040 & .064 & .068 & a_1 \\
 & & .115 & .091 & \underline{.059} & \underline{.071} & .077 & .080 & .074 & .059 & .058 & .081 & a_2 \\
C = & & .121 & .098 & .081 & .117 & .088 & .083 & \underline{.044} & \underline{.037} & .074 & .102 & a_3 \\
 & & .114 & .091 & .078 & .118 & .092 & .095 & .087 & .068 & \underline{.050} & \underline{.064} & a_4 \\
 & & .129 & .103 & .095 & .130 & .101 & .099 & .091 & .068 & .082 & .105 & a_5
\end{array}$$

The elements of the C matrix are

$$c_{ji} = \sum_{n=1}^{n=N} l_{jn}\, q_n\, p_{ni} = \begin{cases} \displaystyle\sum_{\substack{n=1 \\ n \neq j}}^{n=N} q_n\, p_{ni} & \text{for } j = 1, \ldots, J-1 = 4 \\[2ex] \displaystyle\sum_{n=1}^{n=N} q_n\, p_{ni} & \text{for } j = J = 5 \end{cases}$$

If we multiply c_{ji} $(j \leq 4)$ by $1 = (c_{5i})/(c_{5i})$ we have

$$c_{ji} = c_{5i} \sum_{\substack{n=1 \\ n \neq j}}^{n=N} (q_n\, p_{ni}/c_{5i}) = c_{5i} \sum_{\substack{n=1 \\ n \neq j}}^{n=N} \text{Prob}\,(\theta_n \mid x_i) \text{ for } j=1, \ldots, J-1$$

Because θ_j and a_j (or θ_n and a_n) are pairs made up of a state of nature and a decision that this is the true state, it is natural for a given $j \leq 4$ to call

$$\sum_{\substack{n=1 \\ n \neq j}}^{n=N} \text{Prob}\,(\theta_n \mid x_i) = 1 - \text{Prob}\,(\theta_j \mid x_i)$$

the probability of an error (i.e. misclassification) when x_i is observed and a_j is taken. Because c_{5i} must be positive,

selection of the j to be associated with c_i by the criterion $\underset{j}{\text{Min}}\, c_{ji}$ is here equivalent to selection of j so as to minimize the probability of an error given x_i. The "minimum error" strategy which results from the application of this criterion for $i = 1,\ldots,I$ has been put forward in its own right as being useful when one does not wish to estimate a loss function in a classification problem. The discussion in the article by Birnbaum and Maxwell is based on the use of such strategies. For our purpose it suffices to note that we have informally shown a sufficient condition for the optimality of the minimum error strategy, namely that the loss matrix can be put in the form

$l_{nn} = v$ where a_n is the decision that θ_n obtains $(n = 1,\ldots,N)$
$l_{jn} = w > v$ for $j \neq n;\ j,\ n = 1,\ldots,N$
$l_{jn} \geqq w$ for $j = N+1,\ldots,J$ and $n = 1,\ldots,N$

We now turn to an analysis of the changes which take place in the optimal strategy as u varies. In order to obtain u as a fraction of the cost of misclassification directly, we carry out the analysis procedure using the loss matrix

$$L^* = L + u\,E$$

where L is the loss matrix for $u = 0$ and E is a 5x4 matrix with $e_{jn} = 0$ for $j,\ n = 1,\ldots,4$ and $e_{5n} = 1$ for $n = 1,\ldots,4$.

The matrix $C = L\phi P$ used to specify the optimal strategy when $u = 0$ is the same as the C matrix given earlier except that every element in the last row is zero. The numbers in the last row of the earlier C matrix constitute the last row of $D = E\phi P$. All other rows of D are zero. Therefore as u varies only the last row of $C^* = C + uD$ will change. Note that here we use the last row of C^* rather than the first row as was done in the general discussion.

Following the analysis procedure of section A2, we first underscore $c_{j_i i} = \underset{j \le 4}{\text{Min}} \; c_{ji}$ for $i = 1,\ldots,10$. We next compute u_i and determine the superscripts to be attached to u_i, c_{5i}, and $c_{j_i i}$ for $i = 1,\ldots,10$. Then we reorder the columns of C to conform to the ranking of the u_i's. If we wish to express $E(s_o)$ as a piece-wise linear function of u, we may carry out the computational procedure of section B2 in a 5x10 table under the C matrix to obtain the required coefficients using the optimal strategy for $u = 0$, s_{oo}, as the starting point for the computations of $tr(S_{oi}C)$ and $tr(S_{oi}D)$. The information obtained from the two analysis procedures is summarized on the following page.

We now can specify the optimal strategy matrix and compute the corresponding minimum expected loss for any value of u. If the expected cost of the procedure to be used in the continued testing implied by a_5 is less than 23.2% of the cost of misclassification the decision to refuse to classify

$$\begin{array}{cccc}
 & \begin{array}{cccccccccc} (1) & (2) & (3) & (4) & (5) & (6) & (7) & (8) & (9) & (10) \\ x_2 & x_1 & x_6 & x_5 & x_7 & x_8 & x_4 & x_9 & x_{10} & x_3 \end{array} & & \\
 & \left(\begin{array}{cccccccccc}
\underline{.029}^{+} & \underline{.037}^{+} & \underline{.039}^{+} & \underline{.046}^{+} & .068 & .040 & .094 & .064 & .068 & .067 \\
.091 & .115 & .080 & .077 & .074 & .059 & \underline{.071}^{+} & .058 & .081 & \underline{.059}^{+} \\
.098 & .121 & .083 & .083 & \underline{.044}^{+} & \underline{.037}^{+} & .117 & .074 & .102 & .081 \\
.091 & .114 & .095 & .092 & .087 & .068 & .118 & \underline{.050}^{+} & \underline{.064}^{+} & .078 \\
0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++} & 0^{++}
\end{array}\right) & & \begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{array} \\
\begin{array}{c} -\infty \\ \\ \\ 0 \\ 1.003 \end{array} & \begin{array}{cccccccccc}
.282^{++} & .287^{++} & .394^{++} & .455^{++} & .485^{++} & .544^{++} & .546^{++} & .610^{++} & .610^{++} & .621^{++} \\
-.103 & -.129 & -.099 & -.101 & -.091 & -.068 & -.130 & -.082 & -.105 & -.095 \\
.029 & .037 & .039 & .046 & .044 & .037 & .071 & .050 & .064 & .059 \\
.029 & .066 & .105 & .151 & .195 & .232 & .303 & .353 & .417 & .476 \\
.900 & .771 & .672 & .571 & .480 & .412 & .282 & .200 & .095 & 0
\end{array} & \begin{array}{c} \infty \\ \\ \\ \\ \\ \end{array} & \begin{array}{c} u_i \\ g_i \\ h_i \\ tr(S_{oi}C) \\ tr(S_{oi}D) \end{array}
\end{array}$$

pending further tests will always be made, and the minimum expected cost is 1.033 u for u $\leq$ .282. If on the other hand the expected cost of the procedure to be used in the additional tests involved in a_5 is more than 62.1% of the cost of a misclassification, we will always make a classification using the minimum error strategy, and the minimum expected loss is .476 regardless of u $\geq$ .621.

As u decreases below .621 (or increases above .282) we make a classification when some indications are observed but refuse to classify patients on the basis of others. Thus if u = .50 we would classify patients for whom x_1, x_2, x_5, x_6, or x_7 was observed but would refuse to classify otherwise. In this case the minimum expected loss is .195 + .50(.480) = .435.

Such a strategy may be called a "restricted minimum error" strategy because it is to follow the minimum error strategy

subject to the restriction that the probability of a misclassification be less than or equal to u. This becomes apparent when we note that with the given L and W matrices our computation of u_i gives

$$u_i = \frac{\underset{j<5}{\mathrm{Min}}\, c_{ji} - c_{5i}}{d_{5i}} = \underset{j\leq 4}{\mathrm{Min}} \left(\sum_{\substack{n=1 \\ n\neq j}}^{n=4} q_n p_{ni} \Big/ \sum_{n=1}^{n=4} q_n p_{ni} \right) = \underset{j\leq 4}{\mathrm{Min}} \left(\sum_{\substack{n=1 \\ n\neq j}}^{n=4} \mathrm{Prob}(e_n | x_i) \right)$$

which is the minimum probability of misclassification available when x_i is observed. Thus our refusal to classify on the basis of x_3, x_4, x_8, x_9, or x_{10} when u = .50 is equivalent to a refusal to classify on the basis of an indication for which the minimum probability of error is greater than .50.

In calling such a strategy a restricted minimum error strategy, we follow Chow (1957) who considers such strategies in the context of devising circuits in a character recognition system (e.g. the devices used to process machine readable checks in many banks.)

Chapter 5

Sensitivity Analysis Applied to Probability Distributions

In the preceding chapters our discussion of the methods of sensitivity analysis was oriented toward the resolution of problems which may arise in connection with the loss matrix in a finite statistical decision problem. However the methods themselves are independent of our interpretation of the parameters under analysis in a specific problem. The applicability of the methods hinges on the satisfaction of the conditions given in section IV A of Chapter 2. In this chapter we consider how these conditions can be satisfied when the parameters under scrutiny influence one of the probability distributions in a finite statistical decision problem and present two examples to illustrate the procedure.

I. The Criterion Matrix as a Linear Function of the Elements of a Probability Distribution: In order to employ the methods of Chapters 3 and 4 analyze the sensitivity of an optimal strategy for a finite statistical decision problem to variations in M parameters $u_1, \ldots, u_M$ it must be possible to express the criterion matrix as a linear function

$$C = C_0 + \sum_{m=1}^{m=M} u_m C_m$$

of the parameters where C_m is an arbitrary JxI matrix for m = 0, 1,...,M.

In Chapter 1 we discussed two ways in which probability information might be available in a finite statistical decision problem. Below we briefly review this discussion and in each case point out how the criterion matrix can be expressed as a linear function of the elements of one of the relevant probability distributions.

I A. A Prior Distribution over States of Nature and Conditional Distributions Over Indications: We arrange the elements q_n = Prob(θ_n) on the principal diagonal of the matrix $\not{Q}$(NxN) and the elements p_{ni} = Prob($x_i \mid \theta_n$) in the matrix P(NxI), the n-th row of which, P_n'(1xI), gives the probability distribution over indications conditional on θ_n being the true state of nature. Then we express the criterion matrix as

$$C = L\not{Q}P$$

where L(JxN) is the loss matrix, the n-th column of which, L_n(Jx1), contains the losses incurred when actions are taken in the presence of θ_n.

To express the criterion matrix as a linear function of the elements of the prior distribution over states of nature we form N JxI matrices

$$C_n = L_n P_n' \qquad n = 1,\ldots,N$$

and write the criterion matrix as

$$C = \sum_{n=1}^{n=N} q_n C_n$$

To express the criterion matrix as a linear function of the elements of the conditional distributions over indications, we let C_{ni} (JxI) be a matrix having the n-th column of $L\mathcal{Q}$ (which is $q_n L_n$) as its i-th column and zeros elsewhere for $i = 1,\ldots,I$ and $n = 1,\ldots,N$. Then we write the criterion matrix as

$$C = \sum_{n=1}^{n=N} \sum_{i=1}^{i=I} p_{ni} C_{ni}$$

If we are interested only in studying the elements of the conditional distribution for a particular state of nature, say θ_N, we let

$$C_o = \sum_{n=1}^{N-1} \sum_{i=1}^{i=I} p_{ni} C_{ni}$$

and express the criterion matrix as

$$C = C_o + \sum_{i=1}^{i=I} p_{Ni} C_{Ni}$$

I B. A Prior Distribution Over Indications and Posterior Distributions over States of Nature: We arrange the elements $p_i = \text{Prob}(x_i)$ along the principal diagonal of the matrix $\mathcal{P}$(IxI) and the elements $q_{ni} = \text{Prob}(\theta_n \mid x_i)$ in the matrix Q(NxI). Then we express the criterion matrix as either

$$C = LQ\mathcal{P} \qquad \text{or} \qquad C = LQ$$

depending on whether we wish to compute the minimum expected loss or simply determine the optimal strategy. Because the optimal strategy is totally insensitive to variations in the elements of a prior distribution over indications (section II of Chapter 1), we will not consider the criterion matrix as a linear function of the elements of that distribution in this chapter.

To express the criterion matrix as a linear function of the elements of the posterior distributions we let $C_{ni}(J \times I)$ be a matrix having zeros everywhere except its i-th column which is L_n or $p_i L_n$ (according to whether we only wish to investigate changes in the optimal strategy or also wish to determine the minimum expected loss) for $n = 1, \ldots, N$ and $i = 1, \ldots, I$. With this notation we can write the criterion matrix as

$$C = \sum_{i=1}^{i=I} \sum_{n=1}^{n=N} q_{ni} C_{ni}$$

If we are interested only in studying the elements of the posterior distribution for a particular indication, say x_I, we let

$$C_o = \sum_{i=1}^{I-1} \sum_{n=1}^{n=N} q_{ni} C_{ni}$$

and express the criterion matrix as

$$C = C_o + \sum_{n=1}^{n=N} c_{nI} C_{nI}$$

Since the first I-1 columns of C_{nI} $(n = 1,\ldots,N)$ are zero, only the I-th column of this criterion matrix need be considered in the sensitivity analysis.

I C. Applicability of Sensitivity Analysis to Probability Distribution Parameters: In sections A and B above we have shown how the criterion matrix can be expressed as a linear function of the elements of either of the probability distributions involved in the statistical decision problem. Therefore, presuming that only linear constraints which define a bounded set are imposed, we may use the methods of Chapters 3 and 4 to study the sensitivity of an optimal strategy to variations in the elements of one of the distributions or to variations in parameters which influence these elements linearly, i.e. the individual probabilities are linear functions of the parameters.

Questions about probability distributions often take the form of questions about the moments (e.g. means, variances, and measures of skewness) of the distribution or perhaps about the true values of other parameters of fitted theoretical distributions. The individual elements of a discrete probability distribution generally cannot be expressed as linear functions of the moments or other parameters of the distribution, and

$$[illegible]$$

Since the first k-1 columns of C_{kj} (j = 1,...,k) are zero, only the k-th column of this constraint matrix need be considered in the sensitivity analysis.

5.6. Applicability of sensitivity analysis to probability distribution parameters. In sections 4 and 5 above we have shown that the criterion values can be expressed as a linear function of the elements of either of the probability distributions involved in the statistical decision problem. Therefore, assuming that only linear constraints which define a bounded set are imposed, we may use the results of Chapters 3 and 4 to study the sensitivity of an optimal strategy to variations in the elements of [illegible] in the distributions or to variations in parameters which influence these elements linearly, i.e., the individual probabilities are linear functions of the parameters.

Questions about probability distributions often take the form of questions about the moments (e.g., means, variances, and measures of skewness) of the distribution or perhaps about the true values of other parameters of fitted theoretical distributions. The individual moments of a discrete probability distribution generally cannot be expressed as linear functions of the elements or other parameters of the distribution, and therefore the condition that the criterion matrix be expressible as a linear function of the parameters under study limits the applicability of the methods of Chapters 3 and 4 more significantly when the parameters under study influence a probability distribution than when they influence the loss matrix. However many fruitful questions can be phrased in terms of the individual probabilities or in terms of parameters which satisfy the linearity condition. Two such questions are illustrated in the following section.

II. Examples of Sensitivity Analysis: To illustrate the application of the methods of sensitivity analysis to probability distribution parameters, we will present two examples based on the following classification problem. Overall and Williams (1963) discuss the use of a minimax error strategy for the diagnosis of hypofunction or hyperfunction of the thyroid gland. By means of a follow-up study of patients after they had been under treatment for at least six months, Overall and Williams classified each patient in their sample as belonging in one of the categories (our states of nature) listed below.

State of Nature: the patient belongs in category	Prob (θ_n) q_n
θ_1: Hypothyroid	.15
θ_2: Euthyroid	.65
θ_3: Hyperthyroid	.20
	1.00

Overall and Williams investigated twenty-one indices of thyroid functioning individually and in various combinations. Here we will consider only the index, "6 hour thyroid uptake of iodine 131," which Overall and Williams found to be the best single indicator among the 21 they studied. The prior distribution above and the following conditional probability distribution matrix are taken from the article by Overall and Williams.

Indications

6 hour iodine 131 uptake (%)

$$\begin{array}{ccccccc}
\text{Below } 2 & 2-7 & 8-27 & 28-35 & \text{Above } 35 & & \\
x_1 & x_2 & x_3 & x_4 & x_5 & & \\
.33 & .57 & .10 & 0 & 0 & \theta_1 & \text{States of Nature} \\
.01 & .10 & .85 & .03 & .01 & \theta_2 & \\
0 & .01 & .04 & .16 & .79 & \theta_3 &
\end{array}$$

After arranging the prior distribution along the principal diagonal of a matrix, $\emptyset$, we may compute the joint probability distribution matrix.

$$\emptyset P = \begin{pmatrix} .050 & .086 & .015 & 0 & 0 \\ .007 & .065 & .553 & .020 & .007 \\ 0 & .002 & .008 & .032 & .158 \end{pmatrix}$$

We consider three actions, i.e. classifying the patient in one of the three categories. In this problem it will be convenient to use a loss matrix of the following form.

$$\begin{array}{cc} & \text{States of Nature} \\ & \begin{array}{ccc} \theta_1 & \theta_2 & \theta_3 \end{array} \\ L = & \left(\begin{array}{ccc} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{array} \right) \end{array} \quad \begin{array}{l} \text{Actions} \\ a_1\text{: classify hypothyroid} \\ a_2\text{: classify euthyroid} \\ a_3\text{: classify hyperthyroid} \end{array}$$

As was noted in connection with example 4 of Chapter 4, where we used a loss matrix having zeros on its principal diagonal and ones elsewhere, such a loss matrix will lead to a minimum error strategy. With this loss matrix we may proceed to compute the criterion matrix.

$$\begin{array}{cc} & \text{Indications} \\ & \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ C = & \left(\begin{array}{ccccc} \underline{-.050} & \underline{-.086} & -.015 & 0 & 0 \\ -.007 & -.065 & \underline{-.553} & -.020 & -.007 \\ 0 & -.002 & -.008 & \underline{-.032} & \underline{-.152} \end{array} \right) \end{array} \begin{array}{l} \\ \\ a_1 \\ a_2 \quad \text{Actions} \\ a_3 \end{array}$$

We have underscored $c_i = \underset{j}{\text{Min}}\, c_{ji}$ for $i = 1, \ldots, 5$ in the criterion matrix, and can now construct an optimal strategy matrix or can describe the optimal strategy more compactly by the quintuplet (1 1 2 3 3).

II A. Sensitivity Analysis for Elements of a Prior Distribution Over States of Nature: In this section we will study the sensitivity of the optimal strategy to variations in the elements of the prior probability distribution over states of nature.

II A 1. Preliminary discussion: Before examining the thyroid problem, some comments on the possible reasons for such a study seem to be in order. We first note that in general if we restrict the elements, q_n, only to be non-negative and to sum to one, application of the methods of sensitivity analysis will produce the set of "admissible" strategies. A pure strategy, s_k, is "admissible" if there does not exist another strategy, $s_{k'}$ (pure or mixed), such that

$$Y_k - Y_{k'} \overset{>}{(\geq)} 0$$

with the strict inequality holding for at least one element. In Chapter 1 we defined $Y_k(N\times 1)$ to be the vector of the expected losses for strategy s_k conditional on θ_n $(n = 1,\ldots,N)$ being the true state of nature and defined a pure strategy to be one of the $K = J^I$ strategies having strategy matrices of the form described there. For this definition we may think of a "mixed" strategy (we will not use the concept otherwise) as being a procedure for selecting among two or more specified pure strategies by means of a random device, e.g. flipping a coin.

Admissible strategies have attracted considerable attention in statistical decision theory because an optimal strategy may always be found among the admissible strategies. This is true regardless of the prior distribution over states of nature when "optimality" is defined as in Chapter 1. It is also true in decision theory under "uncertainty" (i.e. cases where a prior

distribution in a relative frequency sense is either not available or not meaningful) when we adopt one of the usual optimality criteria, Minimax, Minimax Risk, etc. Therefore we need never investigate any but admissible strategies while searching for an optimal strategy.

To exemplify how one might go about such an investigation, we note that a part of what Birnbaum and Maxwell (1960) propose can be described as follows. Instead of attempting to evaluate the losses associated with each (a_j, θ_n) pair in a classification problem, we use a loss matrix such as the one given above for the thyroid problem and we examine successively some or all admissible strategies, regarding the prior distribution over states of nature merely as an hypothetical construct. For each admissible strategy, s_a, we evaluate the $N(N-1)$ probabilities of classifying an individual from θ_n as being in θ_j (error probabilities)

$$\text{Prob } (a_j \mid \theta_n, s_a) = P_n' S_{aj} \qquad n = 1,\ldots,N \; ; \; n \neq j = 1,\ldots,N$$

and the N probabilities of classifying an individual from θ_n as being in θ_n (probabilities of correct classification)

$$\text{Prob } (a_n \mid \theta_n , s_a) = P_n' S_{an} \qquad n = 1,\ldots,N$$

where S_{aj} is the j-th column of the strategy matrix S_a and a_j is interpreted as a classification of θ_j for $j = 1,\ldots,N$. If

some of the N^2 probabilities in the matrix $P\Theta_a$ are unsatisfactory, we make an alteration in the prior distribution to obtain another admissible strategy and carry out a similar evaluation. The process is repeated until a strategy is discovered which produces a satisfactory pattern of "error probabilities" and "probabilities of correct classification."

Whether or not such a procedure is to be preferred to an explicit evaluation (albeit imprecise) of the losses associated with each (a_j, θ_n) pair will depend on the particular problem. However if we elect to use a procedure which, like this one, involves studying some properties of a number of admissible strategies, it seems desirable to have a systematic method for producing some or all of the admissible strategies.

A second general reason for interest in a sensitivity analysis applied to variations in the elements of the prior distribution over states of nature is related to the sources of our information about the various components of the statistical decision problem. Often the conditional distributions, P, and the loss matrix, L, can be determined under controlled conditions without reference to the prior distribution over states of nature. However the derived strategy will be employed in many different environments each of which will have its peculiar prior distribution. Then it seems desirable to have optimal strategies catalogued as a function of the prior distribution over a fairly wide range in order to be

able to apply the procedure without a complete recomputation in each environment and/or to be able to alter the procedure readily as more data is gathered for improved estimation of the prior distribution in each environment.

II A 2. Sensitivity analysis for variations in elements of thyroid problem prior distribution: Returning now to the thyroid problem, the first step in preparing for an analysis of the sensitivity of the optimal strategy to variations in the prior distribution is to express the criterion matrix as a linear function

$$C = \sum_{n=1}^{n=3} q_n C_n$$

of the prior probabilities. With the given loss matrix the n-th row of the matrix C_n is simply the negative of the n-th row of the conditional distribution matrix and all other rows of C_n are zero for n = 1, 2, 3. For example

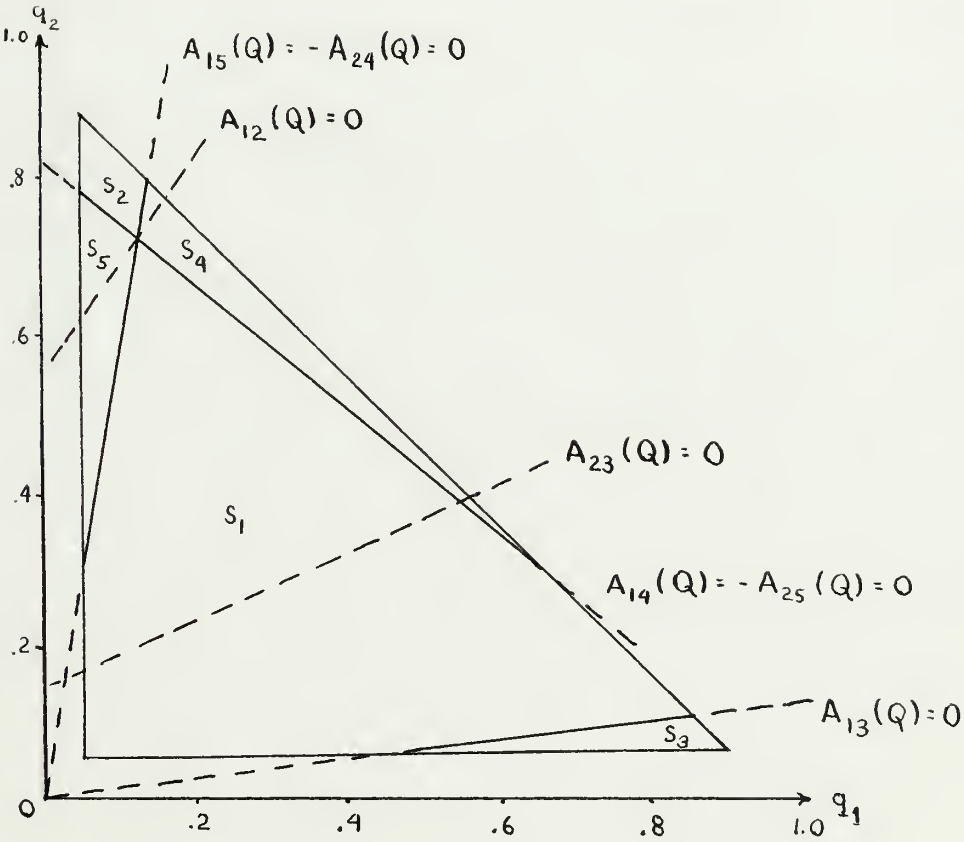
$$C_2 = L_2 P_2' = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} (.01 \ .10 \ .85 \ .03 \ .01) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -.01 & -.10 & -.85 & -.03 & -.01 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

In this example we will determine a subset of the set of admissible strategies by permitting the elements of $Q' = (q_1, q_2, q_3)$ to very subject to the linear constraints

$$q_n \geq .05 \quad \text{for } n = 1, 2, 3$$

$$\sum_{n=1}^{n=3} q_n = 1 \qquad \text{or} \quad q_3 = 1 - (q_1 + q_2)$$

The equation can be used to reduce the problem to a two dimensional one which can be studied graphically. Below we picture the regions in the $(q_1\ q_2)$ plane over which particular strategies are optimal.


q_2
1.0
.8
.6
.4
.2
0
.2
.4
.6
.8
1.0
q_1
$A_{15}(Q) = -A_{24}(Q) = 0$
$A_{12}(Q) = 0$
S_2
S_5
S_4
$A_{23}(Q) = 0$
S_1
$A_{14}(Q) = -A_{25}(Q) = 0$
$A_{13}(Q) = 0$
S_3

The five optimality regions given above may be determined as follows. We compute the optimal strategy at each vertex of the triangle defined by

$$
\begin{array}{lrcl}
(1) & q_1 & - \; .05 & \geq 0 \\
(2) & q_2 & - \; .05 & \geq 0 \\
(3) & -(q_1+q_2) & + \; .95 & \geq 0
\end{array}
$$

and find the following strategies to be optimal.

Optimal Strategies	$C = \sum_{n=1}^{n=3} q_n C_n$	Point $Q=(q_1 q_2 q_3)$	Constraints Holding as Equalities
s_1: (11233)	$\begin{pmatrix} \underline{-.017} & \underline{-.029} & -.005 & 0 & 0 \\ -.001 & -.005 & \underline{-.043} & -.002 & -.001 \\ 0 & -.009 & -.036 & \underline{-.144} & \underline{-.711} \end{pmatrix}$	(.05 .05 .9)	(1)(2)
s_2: (12223)	$\begin{pmatrix} \underline{-.017} & -.029 & -.005 & 0 & 0 \\ -.009 & \underline{-.090} & \underline{-.765} & \underline{-.027} & -.009 \\ 0 & -.001 & -.002 & -.008 & \underline{-.040} \end{pmatrix}$	(.05 .9 .05)	(1)(3)
s_3: (11133)	$\begin{pmatrix} \underline{-.297} & \underline{-.513} & \underline{-.090} & 0 & 0 \\ -.001 & -.005 & -.043 & -.002 & -.001 \\ 0 & -.001 & -.002 & \underline{-.008} & \underline{-.040} \end{pmatrix}$	(.9 .05 .05)	(2)(3)

We then determine the lines, $A_{vw}(Q) = E(s_v) - E(s_w) = 0$, on which the expected losses of each pair of strategies are equal. The equations of these three lines are

$$
\begin{aligned}
A_{12}(Q) &= -.41\,q_1 + .29\,q_2 - .16 = 0 \\
A_{13}(Q) &= .10\,q_1 - .85\,q_2 = 0 \\
A_{23}(Q) &= .51\,q_1 - 1.14\,q_2 + .16 = 0
\end{aligned}
$$

Upon examining the intersections of $A_{13}(Q) = 0$ with the sides of the triangle we find the following two points

Optimal Strategies	$C = \sum_{n=1}^{n=3} q_n C_n$	Point $Q' = (q_1 q_2 q_3)$	Constraints Holding as Equalities
s_1, s_3	$\begin{pmatrix} \underline{-.180} & \underline{-.242} & \underline{-.043} & 0 & 0 \\ -.001 & -.005 & \underline{-.043} & -.002 & -.001 \\ 0 & -.005 & -.021 & \underline{-.024} & \underline{-.415} \end{pmatrix}$	(.43 .05 .52)	(2) A_{13}
s_1, s_3	$\begin{pmatrix} \underline{-.221} & \underline{-.425} & \underline{-.085} & 0 & 0 \\ -.001 & -.010 & \underline{-.085} & -.003 & -.001 \\ 0 & -.001 & -.002 & \underline{-.008} & \underline{-.040} \end{pmatrix}$	(.85 .1 .05)	(3) A_{13}

Therefore we can conclude that s_3 is optimal over the triangle having those two points and (.9 .05 .05) as vertices.

Furthermore upon examining the intersections of $A_{12}(Q) = 0$ with the sides of the triangle we discover two more strategies which are optimal in the convex set defined by (1), (2), and (3). These two strategies are s_4, (11223), and s_5, (12233). Their discovery requires us to expand our list of lines $A_{vw}(Q) = 0$ to include the following equations.

$$
\begin{aligned}
A_{14}(Q) &= .16\,q_1 + .19\,q_2 - .16 = 0 \\
A_{15}(Q) &= -.57\,q_1 + .10\,q_2 = 0 \\
A_{24}(Q) &= .57\,q_1 - .10\,q_2 = 0 \\
A_{25}(Q) &= -.16\,q_1 - .19\,q_2 + .16 = 0 \\
A_{34}(Q) &= .06\,q_1 + 1.04\,q_2 - .16 = 0 \\
A_{35}(Q) &= -.67\,q_1 + .95\,q_2 = 0 \\
A_{45}(Q) &= -.73\,q_1 - .09\,q_2 + .16 = 0
\end{aligned}
$$

We note that $A_{14}(Q) = -A_{25}(Q) = 0$ and $A_{15}(Q) = -A_{24}(Q) = 0$. If we plot both of these lines and determine the optimal strategies

at the points where each of them cuts a side of the triangle or intersects $A_{12}(Q) = 0$, we can tabulate the following results.

Optimal Strategies	$C = \sum_{n=1}^{3} q_n C_n$	Point $Q' = (q_1 q_2 q_3)$	Constraints Holding as Equalities
s_1, s_5	$\begin{pmatrix} \underline{-.017} & \underline{-.029} & -.005 & 0 & 0 \\ -.003 & \underline{-.029} & \underline{-.342} & -.009 & -.003 \\ 0 & -.007 & -.027 & \underline{-.106} & \underline{-.525} \end{pmatrix}$	(.05 .29 .66)	(1) A_{15}
s_2, s_5	$\begin{pmatrix} \underline{-.017} & -.029 & -.005 & 0 & 0 \\ -.008 & \underline{-.080} & \underline{-.690} & \underline{-.024} & -.008 \\ 0 & -.002 & -.006 & \underline{-.024} & \underline{-.119} \end{pmatrix}$	(.05 .80 .15)	(1) A_{14}
s_1, s_2 s_4, s_5	$\begin{pmatrix} \underline{-.042} & \underline{-.073} & -.013 & 0 & 0 \\ -.007 & \underline{-.073} & \underline{-.624} & \underline{-.022} & -.007 \\ 0 & -.001 & -.006 & \underline{-.022} & \underline{-.109} \end{pmatrix}$	(.13 .73 .14)	A_{12} A_{14} A_{15}
s_2, s_4	$\begin{pmatrix} \underline{-.042} & \underline{-.081} & -.014 & 0 & 0 \\ -.008 & \underline{-.081} & \underline{-.687} & \underline{-.024} & -.008 \\ 0 & -.001 & -.002 & -.008 & \underline{-.040} \end{pmatrix}$	(.14 .81 .05)	(3) A_{15}
s_1, s_4	$\begin{pmatrix} \underline{-.226} & \underline{-.390} & -.068 & 0 & 0 \\ -.003 & -.027 & \underline{-.227} & \underline{-.008} & -.003 \\ 0 & -.001 & -.002 & \underline{-.008} & \underline{-.040} \end{pmatrix}$	(.68 .27 .05)	(3) A_{14}

These five points when added to the five points already determined are sufficient to define the five optimality subsets pictured on the graph. Thus in the graph we have a description of the optimal strategy as a function of the prior distribution over the states of nature to use as we see fit.

II B. Sensitivity Analysis for Elements of Conditional Distributions over Indications: In this section we consider the thyroid problem with a view to determining if the optimal strategy is sensitive to certain variations in the elements of

the conditional probability matrix P when the prior distribution over states of nature is

$$Q' = (.15 \quad .65 \quad .20)$$

The first step in preparing for such a sensitivity analysis is to express the criterion matrix as a linear function

$$C = \sum_{n=1}^{n=N} \sum_{i=1}^{i=I} p_{ni} C_{ni}$$

of the elements of P. In the thyroid problem the i-th column of C_{ni} (which is $q_n L_n$) is simply $-q_n e_n$ where $e_n'(1x3)$ is a unit vector containing a one as its n-th element. Thus for example we have

$$C_{24} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -.65 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We will seek to discover if the strategy described by the quintuplet (11233) will remain optimal when the elements of P vary by ± 10% subject to the restrictions that

$$\sum_{i=1}^{i=5} p_{ni} = 1 \quad \text{for } n = 1, 2, 3$$

That is we will analyse the sensitivity of the optimal strategy to variations in P subject to the linear restrictions

n = 1	n = 2	n = 3
$.297 \leq p_{11} \leq .363$	$.009 \leq p_{21} \leq .011$	$p_{31} = 0$
$.513 \leq p_{12} \leq .627$	$.090 \leq p_{22} \leq .110$	$.009 \leq p_{32} \leq .011$
$.090 \leq p_{13} \leq .110$	$.765 \leq p_{23} \leq .935$	$.036 \leq p_{33} \leq .044$
$p_{14} = 0$	$.027 \leq p_{24} \leq .033$	$.144 \leq p_{34} \leq .176$
$p_{15} = 0$	$.009 \leq p_{25} \leq .011$	$.711 \leq p_{35} \leq .869$
$\sum_{i=1}^{5} p_{1i} = 1$	$\sum_{i=1}^{5} p_{2i} = 1$	$\sum_{i=1}^{5} p_{3i} = 1$

Because there are no links between P_n' and P_m' $(n \neq m)$ in the system of constraints, we can simplify the problem of specifying the extreme points of convex set within which the elements of P must lie by considering the three groups above individually. We determine the extreme points of the smaller convex sets within which $P_n' = (p_{n1},\dots,p_{n5})$ must lie for n = 1, 2, 3 one at a time, and the extreme points of the larger convex set within which the elements of P lie are the combinations obtained by taking any one extreme point from each of the three smaller sets. Furthermore because C_{ni} has a non-zero entry only in cell (ni), once we have determined an extreme point of the convex set within which P_n' lies, we can compute the n-th row of a criterion matrix that corresponds to any extreme point of the set in which P must lie having the given P_n' co-ordinates by

$$\sum_{i=1}^{5} p_{ni} C_{ni} = -q_n P_n'$$

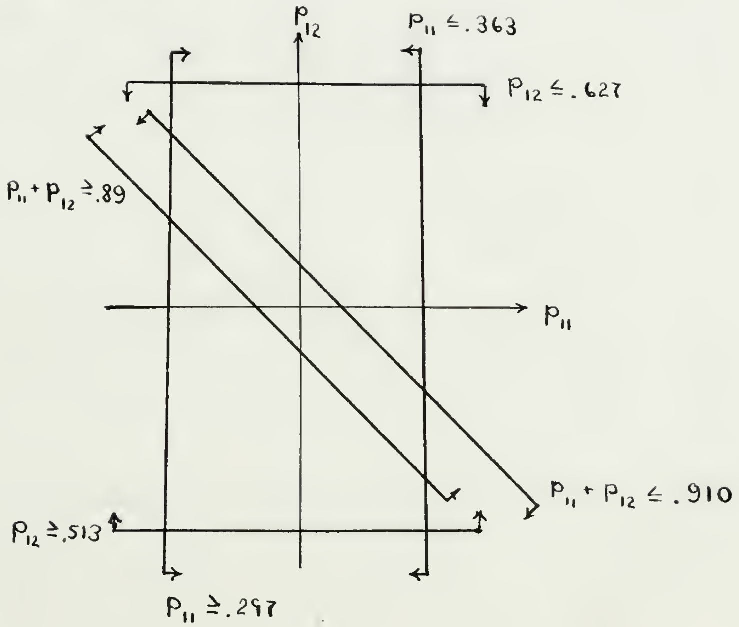
We can specify the extreme points of the convex set within which P_1' lies readily by rewriting the system of constraints as

$.297 \leq p_{11} \leq .363 \qquad p_{13} = 1 - p_{11} - p_{12}$

$.513 \leq p_{12} \leq .627 \qquad p_{14} = 0$

$.890 \leq p_{11} + p_{12} \leq .910 \qquad p_{15} = 0$

The six inequalities define the parallelogram shown below in the $(p_{11}\ p_{12})$ plane.


p_{12}
$p_{11} \leq .363$
$p_{12} \leq .627$
$p_{11} + p_{12} \geq .89$
p_{11}
$p_{11} + p_{12} \leq .910$
$p_{12} \geq .513$
$p_{11} \geq .297$

Instead of tabulating the extreme points themselves, we give $-q_1P_1' = -.15P_1'$ which are the first rows of C matrices corresponding to extreme points having the particular $(p_{11},\dots,p_{15})$

co-ordinates. At the foot of the table we note the maximum and minimum element in each column for future reference.

First row of C under indications

	x_1	x_2	x_3	x_4	x_5
	-.045	-.089	-.017	0	0
	-.045	-.092	-.014	0	0
	-.054	-.079	-.017	0	0
	-.054	-.082	-.014	0	0
Max	-.045	-.079	-.014	0	0
Min	-.054	-.092	-.017	0	0

We can specify the extreme points of the convex set within which P_3' lies most easily by rewriting the system of constraints on P_3' as

$$\begin{array}{ll} p_{31} = 0 & .009 \leq p_{32} \leq .011 \\ p_{35} = 1 - p_{32} - p_{33} - p_{34} & .036 \leq p_{33} \leq .044 \\ .131 \leq p_{32} + p_{33} + p_{34} \leq .209 & .144 \leq p_{34} \leq .176 \end{array}$$

The eight inequalities define a right parallelepiped in the $(p_{32}p_{33}p_{34})$ hyperplane (the constraints on $p_{32} + p_{33} + p_{34}$ turn out to be redundant), and therefore to determine the eight extreme points of the set, we may let p_{32}, p_{33}, and p_{34} take on their extreme values independently and compute $p_{35} = 1 - p_{32} - p_{33} - p_{34}$. Instead of tabulating the extreme

points, we give $-q_3P_3' = -.20P_3'$ which are the third rows of C matrices corresponding to extreme points having the particular $(p_{31},\dots,p_{35})$ co-ordinates.

Third row of C under indications

	x_1	x_2	x_3	x_4	x_5
	0	-.0018	-.007	-.029	-.162
	0	-.0018	-.007	-.035	-.156
	0	-.0018	-.009	-.029	-.161
	0	-.0018	-.009	-.035	-.154
	0	-.0022	-.007	-.029	-.162
	0	-.0022	-.007	-.035	-.155
	0	-.0022	-.009	-.029	-.160
	0	-.0022	-.009	-.035	-.154
Max	0	-.0018	-.007	-.029	-.154
Min	0	-.0022	-.009	-.035	-.162

Specifying the extreme points of the convex set within which P_2' must lie is simply an extension to 4 dimensions of the procedure just used in three dimensions. We let p_{21}, p_{22}, p_{24}, and p_{25} take on their extreme values independently and compute $p_{23} = 1 - p_{21} - p_{22} - p_{24} - p_{25}$, checking each time to insure that the value so obtained satisfies $.765 \leq p_{23} \leq .935$. In this way we specify the sixteen extreme points of a four dimensional orthotope, and for each extreme point we

compute $-q_2 p_2' = -.65 p_2'$ which is the second row of a criterion matrix corresponding to an extreme point having those $(p_{21},\dots,p_{25})$ co-ordinates.

	Second row of U under indications				
	x_1	x_2	x_3	x_4	x_5
	-.006	-.059	-.562	-.018	-.006
	-.006	-.059	-.561	-.018	-.007
	-.006	-.059	-.558	-.021	-.006
	-.006	-.059	-.557	-.021	-.007
	-.006	-.072	-.549	-.018	-.006
	-.006	-.072	-.548	-.018	-.007
	-.006	-.072	-.545	-.021	-.006
	-.006	-.072	-.544	-.021	-.007
	-.007	-.059	-.561	-.018	-.006
	-.007	-.059	-.560	-.018	-.007
	-.007	-.059	-.557	-.021	-.006
	-.007	-.059	-.556	-.021	-.007
	-.007	-.072	-.548	-.018	-.006
	-.007	-.072	-.547	-.018	-.007
	-.007	-.072	-.544	-.021	-.006
	-.007	-.072	-.543	-.021	-.007
Max	-.006	-.059	-.543	-.018	-.006
Min	-.007	-.072	-.562	-.021	-.007

Because the convex sets within which P'_n must lie for $n = 1, 2, 3$ have 4, 16, and 8 extreme points respectively, the union of these three non-intersecting convex sets (the set within which the elements of P lie) has $4(8)(16) = 512$ extreme points. Fortunately it is not necessary to examine each extreme point in order to determine if the strategy (11233) will remain optimal for every P satisfying the given constraints. Because there is no interaction between P'_n and P'_m $(n \neq m)$ in the system of constraints and because P'_n influences only the n-th row of the criterion matrix, we can answer our question with the aid of the following table.

Extreme Values of Elements of C

		Indications					
		x_1	x_2	x_3	x_4	x_5	
Actions	a_1	-.045	-.079	-.014	0	0	Max
		-.054	-.092	-.017	0	0	Min
	a_2	-.006	-.059	-.543	-.018	-.006	Max
		-.007	-.072	-.563	-.022	-.007	Min
	a_3	0	-.002	-.007	-.029	-.154	Max
		0	-.002	-.009	-.035	-.162	Min

This table summarizes the information obtained previously on the maximum and minimum of each element of the criterion matrix as the elements of P vary over the convex set within which they are

constrained. Since the maximums in row 1 are less than the minimums in rows 2 and 3 of columns 1 and 2, a_1 will be taken when x_1 or x_2 is observed. Similarly a_3 will be taken when x_4 or x_5 is observed because the maximums in row 3 are less than the minimums in rows 1 and 2 of columns 4 and 5. Because in column 3 the maximum in row 2 is less than the minimums in rows 1 and 3, a_2 will be taken when x_3 is observed. Therefore the optimality of the strategy described by (11233) is not altered by minor variations in the conditional probability distributions of the type we have considered. The same conclusion would follow from an examination of all 512 extreme points of the set because the element in cell (ji) of the criterion matrix corresponding to any extreme point will be a number which is bracketed by the two numbers in cell (ji) of the table above for all i and j.

In summary we note that in this example sensitivity analysis was relatively easy because of some special properties of the problem studied. On the one hand specification of the co-ordinates of the extreme points of the set was made easier by the absence of links between P'_n and P'_m $(n \neq m)$ in the system of constraints; this is an intrinsic property of the problem. On the other hand we could avoid examining all 512 extreme points because P'_n influences only the n-th row of the criterion matrix. This property of the problem results from our use of the option to add an arbitrary

constant to every element of the loss matrix in order to put it in a form which made sensitivity analysis easy. Had we used a loss matrix such as the one given in example 4 of Chapter 4, P'_n would influence every row of the criterion matrix except the n-th row and sensitivity analysis would have been much more difficult.

Each finite statistical decision problem will have distinctive features and will be flexible enough to allow some latitude in its manner of formulation. Imaginative use of the options available in problem formulation and inventive adaption of the sensitivity analysis procedures to the peculiarities of the problem can make sensitivity analysis easier and perhaps can make the difference between a sensitivity analysis task that is possible and one that is computationally infeasible.

Appendix

Development of a Method for the Investigation of a Convex Set Defined by Linear Constraints

This appendix is devoted to the development of a procedure for investigating a convex set defined by linear constraints with the view of

1. Establishing the consistency or inconsistency of the system of linear constraints which defines the set.
2. Determining if the convex set is unbounded.
3. Revealing the presence of a system of equality constraints if such a system is incorporated in a system of inequality constraints.
4. Identifying those constraints in the system which are redundant in that they are not necessary to the specification of the convex set.
5. Generating the extreme points of the convex set or more generally specifying the set of points which is sufficient to define the convex set.

The desire to achieve these particular ends arose in the context of an analysis of the sensitivity of the optimal strategy of a finite statistical decision problem to variations in several parameters which are subject to linear

constraints. However the method for investigation may be divorced from the specific problem that created a need for it and examined in its own right as is done below.

I. Reduction of a System of both Equalities and Inequalities to one Involving only Inequalities.

I A. Description of Constraints: We begin with a mixed (i.e. both equations and inequalities) system which can be expressed

$$\begin{pmatrix} A_1^\circ \\ \cdots \\ A_2^\circ \end{pmatrix}\begin{pmatrix} U_1^\circ \\ \cdots \\ U_2^\circ \end{pmatrix} + \begin{pmatrix} B_1^\circ \\ \cdots \\ B_2^\circ \end{pmatrix} \begin{matrix} \overset{>}{(\geq)} \\ \cdots \\ = \end{matrix} \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix}$$

where U_1° (Mx1, i.e. M rows and 1 column) and U_2° (Px1) are variable vectors, A_1° (Rx(M+P)) and A_2° (Px(M+P)) are coefficient matrices, and B_1° (Rx1) and B_2° (Px1) are constant vectors. The notation $\overset{>}{(\geq)}$ means that we will consider systems containing both strict and ordinary inequalities as well as systems containing one or the other type exclusively. We usually expect to have $R > P$ and we assume $R \geq M+1$ where M is here taken to be the number of parameters remaining in the system of inequalities after completing the procedure of section I B below. The method is applicable without this assumption but we are primarily interested in sets that are at least potentially bounded. Because the minimal (in the sense of requiring the fewest constraints) bounded convex set defined by linear

inequalities in N-space is a 'simplex' defined by N+1 constraints, a necessary condition for the convex set to be bounded is M $\geq$ N+1.

I B. Reduction Operation: By a series of row operations we try to form a PxP identity matrix in A_2^o and a zero matrix in the corresponding columns of A_1^o. This is essentially the Gaussian reduction technique and can end in two ways.

I B 1. A row of A_2^o has been transformed into a zero row and

a. the corresponding element of B_2^o has been transformed into zero. In this case the constraint is redundant and can be discarded.

b. the corresponding element of B_2^o has been transformed into a nonzero number. In this case the equality constraints are inconsistent among themselves.

I B 2. The desired identity matrix is obtained and

a. a row of A_1^o has been transformed into a row of zeros while the corresponding element of B_1^o has been transformed into a nonpositive (for a strict inequality) or a negative (for an ordinary inequality) number. This is a sufficient condition for the system of equations to be inconsistent with the system of inequalities.

b. The consistency of the combined system remains to be tested (condition a. is not a necessary condition for the two systems to be inconsistent) and we have a result

which, with suitable rearrangement of rows and columns within the two systems, is equivalent to the transformation

$$\begin{pmatrix} I_R & -A^0_{12} A^{0^{-1}}_{22} \\ 0 & A^{0^{-1}}_{22} \end{pmatrix} \begin{pmatrix} A^0_{11} & A^0_{12} \\ A^0_{21} & A^0_{22} \end{pmatrix} \begin{pmatrix} U^0_1 \\ U^0_2 \end{pmatrix} + \begin{pmatrix} I_R & -A^0_{12} A^{0^{-1}}_{22} \\ 0 & A^{0^{-1}}_{22} \end{pmatrix} \begin{pmatrix} B^0_1 \\ B^0_2 \end{pmatrix} \begin{matrix} \overset{>}{(\geq)} & 0 \\ = & 0 \end{matrix}$$

$$\begin{pmatrix} A^0_{11} - A^0_{12}(A^0_{22})^{-1} A^0_{21} & 0 \\ (A^0_{22})^{-1} A^0_{21} & I_P \end{pmatrix} \begin{pmatrix} U^0_1 \\ U^0_2 \end{pmatrix} + \begin{pmatrix} B^0_1 - A^0_{12}(A^0_{22})^{-1} B^0_2 \\ (A^0_{22})^{-1} B^0_2 \end{pmatrix} \begin{matrix} \overset{>}{(\geq)} & 0 \\ = & 0 \end{matrix}$$

where A^0_{11} is RxN and A^0_{22} is (if no equation is redundant) PxP. The transformed constraints can be expressed as

$$U^0_2 = -(A^0_{22})^{-1} \; (B^0_2 + A^0_{21} \; U^0_1)$$

$$(A^0_{11} - A^0_{12} \; (A^0_{22})^{-1} A^0_{21}) \; U^0_1 \; + \; (B_1 - A^0_{12}(A^0_{22})^{-1} \; B^0_2) \; \overset{>}{(\geq)} \; 0$$

The latter relation can be written

$$A \; U + B \; \overset{>}{(\geq)} \; 0$$

where $A = A^0_{11} - A^0_{12} \; (A^0_{22})^{-1} A^0_{21}$ is $R \times N$

$B = B^0_1 - A^0_{12} \; (A^0_{22})^{-1} B^0_2$ is $R \times 1$

$U = U^0_1$ is $N \times 1$

Given a solution to the new system of inequality constraints we can compute U^0_2, or we can use the expression for

U_2^o as a function of U to reduce the dimensionality of the problem that created the system of linear constraints. We note that the series of row operations will not influence the sense of the inequalities because we only add or subtract multiples of equations; we never subtract one inequality from another. This transformation will sometimes reveal some inequality constraint(s) to be redundant by

(i) a row of A_1^o being transformed into a row of zeros while the corresponding element of B_1^o becomes positive (for a strict inequality) or non-negative (for an ordinary inequality). This constraint may be discarded.

(ii) two or more rows of $(A_1^o \;\; B_1^o)$ being transformed into identical rows. All except one of these constraints may be discarded.

(iii) a row of $(A_1^o \;\; B_1^o)$ being transformed into a constraint that is obviously satisfied by every U satisfying some other constraint.

II. Basic Theorems

II A. Theorems on the Consistency of Systems of Inequalities

II A 1. Theorem I: A necessary and sufficient condition that a system of strict inequalities, $AU + B > 0$, be inconsistent is that there exist a vector $K \geq 0$ with at least one $k_r > 0$ $(r = 0, 1, \ldots, R)$ such that $D'K = 0$ where K is $(R+1)\times 1$ and D is $(R+1)\times(N+1)$ given by:

$$D = \left(\begin{array}{c:c} 1 & 0' \\ \hdashline B & A \end{array}\right)$$

This theorem is proved by Carver (1921-22) as Theorem 3 in the form (except for notation):

> "A necessary and sufficient condition that a given system of strict inequalities $A_r(U) > 0$ (where $A_r(U) = \sum_{m=1}^{m=M} a_{rm}u_m + b_r$ for $r = 1,\ldots,R$) be inconsistent is that there should exist a set of $R+1$ constants $k_0, k_1,\ldots,k_R$ such that
>
> $$k_o + \sum_{r=1}^{r=R} k_r A_r(U) \equiv 0$$
>
> with at least one of the k's being positive and none of them being negative."

To demonstrate the equivalence of the two forms we note that

$$k_o + \sum_{r=1}^{r=R} k_r A_r(U) = \begin{array}{l} k_o + \sum_{r=1}^{r=R} k_r \sum_{m=1}^{m=M} a_{rm}u_m + b_r \\ (k_o + \sum_{r=1}^{r=R} k_r b_r) + \sum_{m=1}^{m=M} u_m \sum_{r=1}^{r=R} a_{rm}k_r \end{array}$$

Therefore $k_o + \sum_{r=1}^{r=R} k_r A_r(U) \equiv 0$ if and only if

$$k_o + \sum_{r=1}^{r=R} k_r b_r = 0$$

and

$$\sum_{r=1}^{r=R} a_{rm}k_r = 0 \quad \text{for } m = 1,\ldots,M$$

$D'K = 0$ is a matrix summary of these $R+1$ equations.

II A 2. Theorem II: Given that there exists a $K \geq 0$ with some of its elements positive (say, for notational convenience, that $k_1,\ldots,k_P$ ($P \leq R$) are positive) such that $D'K = 0$,

(1) Then the system of ordinary inequalities $AU + B \geq 0$ is consistent if and only if there exists a U such that:

$$\text{(a)} \quad A_1\,(U) = \cdots = A_P\,(U) = 0$$

$$\text{(b)} \quad A_{P+1}\,(U) \geq 0\ ,\ \ldots\ ,\ A_R\,(U) \geq 0$$

(2) and if the system $AU + B \geq 0$ is consistent it admits only to solutions for which (1)(a) holds.

Proof:

Sufficiency for (1): Clearly if there exists a U such that (a) and (b) hold the system $AU + B \geq 0$ is consistent.

Necessity for (1) and (2): Given that $AU + B \geq 0$ is consistent we have that there exists a U such that $A_r(U) \geq 0$ for $r = 1,\ldots,R$. Now we employ proof by contradiction, and assume that there exists a U and a p between 1 and P such that $A_p(U) > 0$. Then we have

$$k_o + k_p\,A_p(U) + \sum_{\substack{r=1\\ r\neq p}}^{r=R} k_r A_r(U) = 0$$

with $k_p > 0$, $k_r \geq 0$ ($r \neq p$), $k_o \geq 0$, $A_p(U) > 0$, and $A_r(U) \geq 0$ for $r \neq p$ which is to say that a positive number plus a sum

of non-negative numbers is equal to zero which is impossible. Therefore we conclude that there cannot exist a U such that the constraints corresponding to the positive elements of Y hold as strict inequalities.

We note relative to this theorem that if a system of strict inequalities is inconsistent the corresponding system of ordinary inequalities can be consistent only if there are equality constraints incorporated in the inequality constraints; that is, one or more of the inequalities imply relations of the form $L(U) \geq 0$ while others imply $-L(U) \geq 0$ so that the system has solutions only for $L(U)=0$. Furthermore a system of mixed inequalities can be consistent without the corresponding system of strict inequalities being consistent if some of the ordinary inequalities are equivalent to equality constraints which can be satisfied simultaneous with the satisfaction of the strict inequality constraints in the mixed system. This can be seen by partitioning the mixed system into a system of strict inequalities $(A_1U+B_1 > 0)$ and a system of ordinary inequalities $(A_2U + B_2 \geq 0)$. If $A_1U + B_1 > 0$ is inconsistent the mixed system is inconsistent. If $A_1U + D_1 > 0$ and $A_2U + B_2 > 0$ are consistent the mixed system is consistent. If $A_2U + D_2 > 0$ is inconsistent $A_2U + B_2 \geq 0$ is consistent only if it contains hidden equality constraints. These equality constraints may be used to reduce the

dimensionality of U (section I) altering the form of our two systems to $A_1^* U^* + B_1^* > 0$ and $A_2^* U^* + B_2^* \geq 0$ where now $A_2^* U^* + B_2^* > 0$ is assumed to be consistent. If $A_1^* U^* + B_1^* > 0$ is also consistent (i.e. $A_1 U + B_1 > 0$ is consistent and can be satisfied simultaneous with the satisfaction of the equality constraints in $A_2 U + B_2 \geq 0$) the original mixed system is consistent. Otherwise $AU + B \underset{(\geq)}{>} 0$ is inconsistent.

II A 3. Theorem III: There exists a vector $K \geq 0$ with at least one $k_r > 0$ satisfying $D'K = 0$ if and only if the linear programming problem

$$\text{Max } z = (1, d, \ldots, d) K \quad \text{subject to } D'K = 0 \text{ and } K \geq 0$$

where d is a small positive number (i.e. $1 \gg d > 0$)

has an unbounded solution, i.e. an arbitrarily large value for the objective function, z.

Proof:

Sufficiency: Assume we have a basic solution to the linear programming problem with $k_0, k_1, \ldots, k_M$ as basic variables (the argument is unchanged if some other specific k's are basic variables). Each of these variables is zero. If there is an unbounded solution to the linear programming problem it is indicated by the fact that the column(s) of D' associated with some nonbasic variable(s), say k_s, when expressed as a linear combination $(y_{0s}, y_{1s}, \ldots, y_{Ms})$ of the columns of D'

associated with the basic variables contains only nonpositive elements so that:

$$z_s - c_s = y_{0s} + d \sum_{r=1}^{r=M} y_{rs} - d < 0$$

Note that an unbounded optimal solution as described by Hadley (1962) on page 100 is not possible in this problem because d is positive.

Thus we can achieve an arbitrarily large value of the objective function by assigning an arbitrarily large value to k_s while setting $k_r = -y_{rs}k_s \geq 0$ for $r = 1,\ldots,M$, and we will have a vector $K \geq 0$ with at least one $k_r > 0$ satisfying $D'K = 0$.

Necessity: Assume that there exists a $K \geq 0$ with at least one $k_r > 0$ and let g be a positive number. If K satisfies $D'K = 0$ then so will $D'(gK) = 0$. Inserting gK in the objective function of the linear programming problem we have

$$z = g(1,d,\ldots,d)K \quad \text{where } (1,d,\ldots,d)K \text{ is a positive constant}$$

Therefore by making g arbitrarily large we can make the value of the objective function arbitrarily large while maintaining $D'(gK) = 0$ and $gK \geq 0$. Thus we have an unbounded solution.

We note here that the remainder of this appendix will involve linear programming theory and methods. These are

discussed in Hadley (1962) to which we will make frequent reference.

II B. Theorem on Unboundedness (Theorem IV): The convex set defined by a consistent nonhomogeneous system of inequalities, $AU + B \overset{>}{(\geq)} 0$, is unbounded if and only if the corresponding homogeneous system of ordinary inequalities, $AU \geq 0$, admits to a nontrivial solution.

Proof:

Sufficiency: Given that there exists a solution U^{*} with some $u_m^{*} \neq 0$ to $AU \geq 0$ and a solution U^{**} to $AU + B \overset{>}{(\geq)} 0$. Let g be a non-negative scalar and consider $U^{***} = U^{**} + gU^{*}$. This U^{***} is a solution to $AU + B \overset{>}{(\geq)} 0$ for,

$$AU^{***} + B = A(U^{**} + gU^{*}) + B = (AU^{**} + B) + g\,AU^{*}$$

By hypothesis $AU^{**} + B \overset{>}{(\geq)} 0$ and $AU^{*} \geq 0$; therefore with $g > 0$, $AU^{***} + B \overset{>}{(\geq)} 0$. In particular g can be made arbitrarily large so that some element(s) of U^{***} can be made arbitrarily large positive or negative numbers which implies that the nonhomogeneous system defines an unbounded set.

Necessity: If the convex set defined by a consistent nonhomogeneous system is unbounded, three things exist, a solution U^{**}, a vector F having at least one nonzero element, and a non-negative scalar g such that $U^{***} = U^{**} + gF$ is a solution to $AU + B \overset{>}{(\geq)} 0$ for an arbitrarily large g. We have

$$AU^{***} + B = AU^{**} + B + g\,AF \underset{(\geqq)}{>} 0$$

and the inequality holding for arbitrarily large g implies that every element of AF is non-negative. To prove this by contradiction assume that the r-th element of AF which we may denote by

$$s_r = \sum_{m=1}^{m=M} a_{rm} f_m$$

is negative. The corresponding element of $AU^{**} + B$ which we may denote by

$$t_n = \sum_{m=1}^{m=M} a_{rm} u_m^{**} + b_r$$

is non-negative. Now if we take g such that

$$g > \frac{-t_r}{s_r} \geqq 0$$

we will make the r-th element of $AU^{***} + B$ negative, which contradicts the assumption that U^{***} is a solution to the system for arbitrarily large g. Therefore each element of AF must be non-negative, and we can take $U^{*} = F$ as a non-trivial solution to the homogeneous system $AU \geqq 0$.

II C. Theorems on Redundancy of Constraints: A constraint is redundant in a consistent system of M constraints if the statement, "U satisfies the other M-1 constraints' implies, 'U satisfies the constraint in question.' Redundant constraints

are unnecessary to the definition of the convex set in which U is constrained to lie. This section contains four theorems pertaining to the identification of redundant constraints.

II C 1. Theorem V: The constraint $A_s(U) \underset{(\geq)}{>} 0$ is redundant in the consistent system $AU + B \underset{(\geq)}{>} 0$ if and only if there exists a K satisfying $D'K = 0$ with k_s and no other k being negative and at least one k being positive.

Carver (1921-22) proves this in an equivalent form for strict inequality constraints. His proof holds for ordinary inequality constraints. The following theorems will be useful in actually establishing the redundancy or non-redundancy of constraints.

II C 2. Theorem VI: Given a consistent system $AU + B \underset{(\geq)}{>} 0$ where the rank of A is N and this rank is not decreased by the removal of the s-th row of A. The constraint $A_s(U) \underset{(\geq)}{>} 0$ is redundant if and only if there exists a set of N linearly independent vectors $(b_r, a_{r1}, \ldots, a_{rN})$, say $r = 1, \ldots, N < s$ for notational convenience, such that the vector $(b_s, a_{s1}, \ldots, a_{sN})$ can be expressed as a linear combination

$$(b_s, a_{s1}, \ldots, a_{sN}) = l_o\, e_1' + \sum_{r=1}^{r=N} l_r\, (b_r, a_{r1}, \ldots, a_{rN})$$

where e_1' is a 1x(N+1) unit vector (1,0,...,0)

with $l_r \geq 0$ for $r = 0, 1, \ldots, N$ and at least one $l_r > 0$.

Proof: The proof will consist of showing that under the conditions of the first sentence this theorem is equivalent to Theorem V.

Sufficiency: Given l_r is non-negative for $r = 0, 1, \dots, M$ with at least one $l_r > 0$ such that

$$(b_s, a_{s1}, \dots, a_{sM}) = l_o e_1' + \sum_{r=1}^{r=M} l_r (b_r, a_{r1}, \dots, a_{rM})$$

we take $k_s = -1$, $k_r = l_r$ for $r = 0, 1, \dots, M$, $k_r = 0$ for $M < r$, $r \neq s$ and form the vector K required by theorem V.

Necessity: Given a K satisfying $V'K = 0$ with k_s and no other k being negative with at least one $k_r > 0$, we can write directly:

$$k_o e_1' + k_s (b_s, a_{s1}, \dots, a_{sM}) + \sum_{\substack{r=1 \\ r \neq s}}^{r=R} k_r (b_r, a_{r1}, \dots, a_{rM}) = 0$$

or

$$(b_s, a_{s1}, \dots, a_{sM}) = m_o e_1' + \sum_{\substack{r=1 \\ r \neq s}}^{r=R} m_r (b_r, a_{r1}, \dots, a_{rM})$$

where $m_r = - \dfrac{k_r}{k_s}$ for $r = 0, 1, \dots, R$, $r \neq s$; $m_r = 0$ and at least one $m_r > 0$.

For convenience let us renumber so that $s = R$ and the positive m_r's (if any) correspond to $r = 1, \dots, P$. We leave m_o unchanged, and have

$$(b_R, a_{R1}, \dots, a_{RM}) = m_o e_1' + \sum_{r=1}^{r=P} m_r (b_r, a_{r1}, \dots, a_{rM}) + 0$$

a. First assume that only $m_o > 0$. Then we may set $l_o = m_o$. The rank of A is M and this is not decreased by the removal of $(a_{R1},\ldots,a_{RM})$. Therefore among the $(b_r,a_{r1},\ldots,a_{rM})$ for $r = 1,\ldots, R-1$ there are M linearly independent vectors which we may take to be the set required by the theorem, assigning the value zero to their coefficient. Thus we have the linear combination required by the theorem.

b. Now we assume $m_r > 0$ for $r = 1,\ldots,P$ with $P \geq 1$.

b. (i) Assume $(b_1,a_{11},\ldots,a_{1M}),\ldots,(b_P,a_{P1},\ldots,a_{PM})$ are linearly independent. This can occur only if $P \leq M$. In this case we can find M-P vectors which will complete a set of M linearly independent vectors. We then set $l_r = m_r$ for $r = 0, 1,\ldots,P$ and $l_r = 0$ for $r = P+1,\ldots,R-1$, and we have the linear combination required by the theorem.

b. (ii) Assume $(b_1,a_{11},\ldots,a_{1M}),\ldots,(b_P,a_{P1},\ldots,a_{PM})$ are linearly dependent. We can then proceed to reduce the number of positive m_r's until we reach case (i) above as follows. If the vectors are linearly dependent there exists a set $g_1,\ldots,g_P$ not all zero such that:

$$\sum_{r=1}^{r=P} g_r\,(b_r,a_{r1},\ldots,a_{rM}) = 0$$

We select the m_r^* to be made zero by either of the criteria;

$$\frac{m_r^*}{g_r^*} = \underset{r=1,\ldots,P}{\text{Min}} \left(\frac{m_r}{g_r}\,\middle|\, g_r > 0\right) \text{ or } \frac{m_r^*}{g_r^*} = \underset{r=1,\ldots,P}{\text{Max}} \left(\frac{m_r}{g_r}\,\middle|\, g_r < 0\right)$$

and we assign new values to the other m_r's given by:

$$m_r - m_r^* \frac{c_r}{c_r^*}$$

each of which will be non-negative if m_r^* is selected by one of the criteria above. The process is repeated until we have positive coefficients only on a set of linearly independent vectors, which puts us in case (i) above.

This establishes the necessity of the condition.

II C 3. Theorem VII: Given a consistent system $AU + B \underset{(\geq)}{>} 0$ where the rank of A is N and this rank is not decreased by the removal of the R-th row of A. Then $A_R(U) \underset{(\geq)}{>} 0$ is redundant if and only if the linear programming problem:

$$\text{Max } z = (1,d,\ldots,d)\,K \quad \text{subject to } K \geq 0\ ,\ D'K = \begin{pmatrix} b_R \\ a_{R1} \\ \vdots \\ a_{RN} \end{pmatrix}$$

$$\text{where } K' = (k_0, k_1, \ldots, k_{R-1})$$

$$D = \begin{pmatrix} 1 & 0' \\ B_1 & A_1 \end{pmatrix}$$

where A_1 is the matrix of coefficients of the first R-1 constraints and B_1 is the vector of the corresponding constants.

has a basic feasible solution.

A basic feasible solution is one in which only N+1 variables may be positive while the columns of D' corresponding

to these M+1 variables form a basis of (M+1)-space.

Proof: The proof assumes that $(b_r, a_{r1}, \ldots, a_{rM}) \neq 0'$ for $r = 1, \ldots, R$.

Sufficiency: A basic feasible solution to the linear programming problem immediately provides $l_0, l_1, \ldots, l_M$ for Theorem VI.

Necessity: Conversely a set $l_0, l_1, \ldots, l_M$ satisfying Theorem VI constitutes a basic feasible solution to the linear programming problem.

We note that $A_R(U)$ in this theorem can be any constraint for which the conditions of the first sentence hold, not just the last one in some specific listing of the constraints.

II C 4. Theorem VIII: Given M constraints, say $A_1(U), \ldots, A_M(U)$ for convenience, such that $(b_1, a_{11}, \ldots, a_{1M}), \ldots, (b_M, a_{M1}, \ldots, a_{MM})$ are linearly independent.

If there exists a U^* such that:

$$(1) \quad A_1 (U^*) = \cdots = A_M (U^*) = 0$$

while

$$(2) \quad A_{M+1} (U^*) > 0 \text{ , } \ldots \text{ , } A_R (U^*) > 0$$

then the M inequalities $A_1(U), \ldots, A_M(U)$ are not redundant in the consistent system $AU + B \underset{(=)}{>} 0$.

Proof: To prove that one of $A_1(U), \ldots, A_M(U)$, say $A_1(U)$, is

not redundant in the consistent system $AU + B \geq 0$, we exhibit a point U^{**} such that $A_1(U^{**}) < 0$ while $A_r(U^{**}) \geq 0$ for $r = 2,\ldots,n$.

Since $(b_2, a_{21}, \ldots, a_{2M}), \ldots, (b_M, a_{M1}, \ldots, a_{MM})$ are linearly independent, $A_2(U) = \ldots = A_M(U) = 0$ defines a line $L(U)$ in M-space. The intersection of $L(U)$ with the hyperplane $A_1(U) = 0$ defines U^*. (2) above implies that there is a positive distance along $L(U)$ proceeding in either direction from U^* before $L(U)$ intersects one of the hyperplanes $A_{M+1}(U) = 0, \ldots, A_n(U) = 0$. Let the smaller of these two distances be $g > 0$. We may select a point U^{**} on $L(U)$ such that the distance between U^* and U^{**} is less than g and $A_1(U) < 0$. Then U^{**} will satisfy $A_r(U^{**}) \geq 0$ for $r = 2,\ldots,n$. Therefore $A_1(U)$ is not redundant in the consistent system $AU + B \geq 0$.

Clearly a constraint which is not redundant in a consistent system of ordinary inequalities is not redundant in the corresponding consistent system of strict (or strict and ordinary) inequality constraints.

III. The Use of Linear Programming Techniques to Obtain A First Extreme Point Solution, Check Consistency, and Detect Unboundedness.

III A. The Linear Programming Technique: In using the linear programming technique for our problem we will use a tableau of

the following type in which the entry in row m (m = 0,...,M+1) and column r (r = -(P+1),...,-1,0,1,...,M+R) is denoted by y_{mr}.

column

...	-(P+1)	-P	...	-p	...	-1	...	0	1	...	m	...	M	M+1	...	M+r	...	M+R	
								k_o	l_1	...	l_m	...	l_M	k_1	...	k_r	...	k_R	
								1	-W	...	-W	...	-W	d	...	d	...	d	Row
			...		...		0			...		...			...		...		0
			...		...		0			...		...			...		...		1
⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮	...	⋮	...	⋮	⋮	...	⋮	...	⋮	⋮
			...		...		0			...		...			...		...		m
⋮	⋮	⋮	...	⋮	...	⋮	⋮	⋮	⋮	...	⋮	...	⋮	⋮	...	⋮	...	⋮	⋮
			...		...		0			...		...			...		...		M
			...		...		0			...		...			...		...		M+1

The basic variables (i.e. the k's and l's corresponding to the columns which make up the identity matrix in the right hand part of the tableau) are listed in the first column in an order corresponding to the row in which the one appears in their columns. $y_{m,-(P+1)}$ for m = 0,...,M are the "values" (i.e. 1, -W, or d) associated with the variables in the first column. The section made up of $y_{m,-p}$ for m = 0,...,M and -p =

$-1,\ldots,-M$ of the tableau will be used in determining the elements of the vector U_2^0 of section 1. The next column is the "requirement vector" and in our problems will contain only zeros. In fact it may be deleted. The remaining columns (i.e. y_{mr} for $m = 0,\ldots,M+1$; $r = 0,\ldots,M+N$) make up the "body" of the tableau. For any column r the y_{mr}'s ($m = 0,\ldots,M$) express the vector which corresponds to that column in terms of the vectors which make up the basis. In the initial tableau we will have $y_{mr} = 1$ if $m = r$ and $y_{mr} = 0$ if $m \neq r$ for $m = 0,\ldots,M$; $r = 0,\ldots,M$ (i.e. we will start with a $(M+1)x(M+1)$ identity matrix as the left hand side of the body of the tableau), and $y_{0M+r} = b_r, y_{mM+r} = a_{rm}$ ($m = 1,\ldots,M$) for $r = 1,\ldots,N$ (i.e. the right hand side of the body of the tableau is $(B \vdots A)'$. The first $M+1$ elements of the last row of the body of the tableau, $y_{M+1,r}$ for $r = 0,\ldots,M$, will never be used explicitly and so they can be deleted. In linear programming notation we have

$$y_{M+1\ M+r} = \sum_{m=0}^{m=M} y_{m,-(M+1)}\, y_{m\ M+r} - d = z_r - c_r$$

for $r = 1,\ldots,N$

which in the initial tableau is

$$y_{M+1\ M+r} = b_r - V \sum_{m=1}^{m=M} a_{rm} - d \qquad \text{for } r = 1,\ldots,N$$

For future reference we may summarize our initial tableau in matrix notation as

$$\begin{array}{ccccc}
-\ ((A^{o}_{22})^{-1}\, B^{o}_{2})' & 0 & 1 & 0' & D' \\
-\ ((A^{o}_{22})^{-1}\, A^{o}_{21})' & 0 & 0 & I_M & A' \\
* & 0 & 0 & 0' & C'
\end{array}$$

where $((A^{o}_{22})^{-1} B^{o}_{2})'$ is 1xP, $((A^{o}_{22})^{-1} A^{o}_{21})'$ is MxP, the requirement vector is zero, the first vector of the body of the tableau is a unit vector (and since it is in the basis $y_{M+1,0} = 0$), the following M columns contain an MxM identity matrix over and under which are M dimensional zero row vectors, D' is 1xR, A' is MxR, and $C' = D' - \alpha 1'_M A' - d1'_R$ is the 1xR vector of the $(z_r - c_r)$'s.

We are using the linear programming problem

$$\text{Max } z = (1, d, \ldots, d)\, K \quad \text{subject to } H'K = 0\ ,\ K \geq 0$$

as a vehicle for the resolution of our problem, and certain peculiarities of the problem dictate the use of some special linear programming techniques. We have introduced $1_1, \ldots, 1_M$ which in linear programming are called "artificial variables." They are used generally when part of an identity matrix is not available among the vectors of the matrix H', and in our problem we may not have any part of an identity matrix. For purposes of exposition we will assume that we always begin the problem with an identity matrix made up with M artificial vectors. However when some columns of G' form part of an identity matrix, these columns may replace similar "artificial

vectors" in the initial tableau in order to reduce calculations.

The artificial variables are assigned large negative "values" so that if there is a feasible solution (i.e. one with $K \geq 0$) to the maximization problem they will be replaced by "legitimate" variables in the basis. Thus we have $W \gg 1 \gg d > 0$, e.g. $W = 1000$ and $d = 1/1000$. For hand computations W and d need not be assigned explicit values. Noting that at any iteration $y_{M+1, M+r} = v_1(-W) + v_2 1 + v_3 d$ for $r = 1, \ldots, R$ where v_1, v_2, and v_3 are determined by the particular situation, the criterion for the relative magnitude of W, 1, and d is that W is large enough so that if $v_1 \neq 0$, $v_2 1 + v_3 d$ is negligible while d is small enough so that if $v_1 = 0$ and $v_2 \neq 0$, $v_3 d$ is negligible. Contrary to the usual practice in linear programming we will not discard artificial vectors after they are removed from the basis. For additional discussion of the use of artificial variables we refer to Hadley (1962), pp. 116-121, 149.

The fact that our requirement vector is zero means that we will always have a "degenerate solution" of the linear programming problem which creates two difficulties for the standard linear programming simplex technique. Difficulties arise in deciding which vector to remove from the basis at each iteration and "cycling" can occur. These difficulties can be avoided by

using the perturbation technique which in theory consists of replacing the given requirement vector by a perturbed requirement vector

$$b(e) = E_{M+1} + e^{M+1}\,(B \vdots A)'\,E_R, \text{ where here } E_r' = (e, e^2, \ldots, e^r) \text{ for } r \geq 1$$

We refer to Hadley (1962), pp. 175-182, for the proof that the use of this perturbed requirement vector with e a sufficiently small positive number assures a non-degenerate solution to the perturbed linear programming problem thus overcoming the difficulties noted above. In practice it is never necessary to compute b(e) for the following criteria applied at each iteration satisfy our needs.

1. The variable to become a basic variable is k_s, selected by:

$$y_{M+1\,M+s} = \min_{r=1,\ldots,R} (y_{M+1\,M+r} \mid y_{M+1\,M+r} < 0)$$

2. The basic variable to be replaced by k_s is determined by
   a. Variables can be replaced by k_s only if $y_{m,M+s}$ is positive. If this criterion does not uniquely determine the variable to be replaced, go to step b for only those variables such that $y_{m\,M+s}$ is positive.
   b. The minimum of $y_{mo}/y_{m\,M+s}$ over all m carried over from step a. If this minimum is not unique go to step c for only those variables which were tied for minimum here.

c. Repeat this process successively for one column after another, taking those variables which were tied for minimum by the criterion $y_{m\ m-1}/y_{m\ M+s}$ for the calculation of $y_{m\ m}/y_{m\ M+s}$, until a unique minimum is found.

3. Say we decide to replace l_t (or k_t) by k_s from the results of 2. In order to obtain the new tableau, we

a. Replace y_{tr} by y_{tr}/y_{tM+s} for $r = -P, \ldots, -1, 1, \ldots, M+N$. That is we multiply the t-th row by $1/y_{tM+s}$.

b. For $m = 0, \ldots, M+1$; $m \neq t$ replace y_{mr} by $y_{mr} - y_{mM+s}$ (y_{tr}/y_{tM+s}) for $r = -P, \ldots, -1, 0, 1, \ldots, M+N$. That is we subtract from the m-th row $y_{m,\ M+s}$ (the element in the column of k_s) times the new t-th row. Of course this need not be done for y_{M+1r}, $r = -P, \ldots, -1$.

c. Enter the new basic variables in the first column of the new tableau (they are the same as the previous tableau except that l_t or k_t is replaced by k_s) and enter the corresponding "values" in the second column.

An important feature of this procedure as applied to our particular problem is that k_o will remain a basic variable. The criterion for the relative magnitudes of M, l, and d is such that if any $y_{mM+s} > 0$ $(m = 0, \ldots, M)$ there is at least one $y_{mM+s} > 0$ among $m = 1, \ldots, M$ for any k_s selected by 1. And thus the minimum of y_{mo}/y_{mM+s} in 2. will occur

among m = 1,...,M because $y_{00} = 1$ and $y_{m0} = 0$ for m = 1,...,M.

III B. The Results of an Application of the Simplex Algorithm:

III B 1. Summary of possible results: With regard to the presence of artificial vectors (i.e. the columns associated with artificial variables) in the basis and the satisfaction of optimality conditions there are four possible situations.

a. Artificial vectors in the basis and optimality conditions not satisfied.

b. No artificial vectors in the basis and optimality conditions not satisfied.

c. No artificial vectors in the basis and optimality conditions satisfied.

d. Artificial vectors in the basis and optimality conditions satisfied.

We find the first situation at the beginning of the problem and it holds until a series of iterations creates one of the other three situations. We will now consider the significance of each of them in turn.

III B 2. No artificial vectors in the basis and optimality conditions not satisfied: When this situation arises we continue to apply the algorithm until the optimality conditions are satisfied (section 3) or we obtain an indication of an unbounded solution in the form of one or more nonbasic k_r for which $y_{M+1\ N+r} < 0$ and $y_{m,\ N+r} \leq 0$ for m = 0, 1,...,M.

Then by theorems I and III we have a necessary and sufficient condition for the inconsistency of a system of strict inequalities $AU + B > 0$.

If we are working with a system of ordinary inequalities, we still are faced with the possibility suggested by Theorem II that $AU + B \geq 0$ has solutions with "hidden" equality constraints so that $AU + B > 0$ is inconsistent. We turn now to the detection of this situation in the linear programming tableau. For ease of discussion we assume that a gift of prescience permitted the judicious arrangement of the columns of the initial tableau so that $k_o, k_1, \ldots, k_M$ are now the basic variables, so that $k_{M+1}, \ldots, k_T$ $(T \leq R)$ are the variables for which we have an indication of an unbounded solution, and more particularly so that $k_1, \ldots, k_S$ $(S \leq M)$ are the basic variables for which $y_{m,\ M+r} < 0$ $(m = 1, \ldots, S;$ $r = M+1, \ldots, T)$. With this assumption the series of row operations carried out to arrive at the present tableau is equivalent to the following transformation on the body of the tableau.

$$\begin{pmatrix} 1 & -B_1'\,(A_1')^{-1} & 0 \\ 0 & (A_1')^{-1} & 0 \\ 0 & -C_1'\,(A_1')^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0' & B_1' & B_2' & B_3' \\ 0 & I_M & A_1' & A_2' & A_3' \\ 0 & 0' & C_1' & C_2' & C_3' \end{pmatrix} =$$

$$\begin{pmatrix} 1 & -B_1'(A_1')^{-1} & 0' & B_2'-B_1'(A_1')^{-1}A_2' & B_3'-B_1'(A_1')^{-1}A_3' \\ 0 & (A_1')^{-1} & I_M & (A_1')^{-1}A_2' & (A_1')^{-1}A_3' \\ 0 & -C_1'(A_1')^{-1} & 0' & C_2'-C_1'(A_1')^{-1}A_2' & C_3'-C_1'(A_1')^{-1}A_3' \end{pmatrix}$$

where

$C_1' = (B_1' - W1_M' A_1' - d1_M')$ is $1 \times M$ and corresponds to $k_1, \ldots, k_M$

$C_2' = (B_2' - W1_M' A_2' - d1_{T-M}')$ is $1 \times (T-M)$ and corresponds to $k_{M+1}, \ldots, k_T$

$C_3' = (B_3' - W1_M' A_3' - d1_{R-T}')$ is $1 \times (R-T)$ and corresponds to $k_{T+1}, \ldots, k_R$

and where for an indication of an unbounded solution we have:

$$C_2' - C_1' (A_1')^{-1} A_2' < 0$$

$$B_2' - B_1' (A_1')^{-1} A_2' \lneq 0$$

$$(A_1')^{-1} A_2' \lneq 0$$ where the last (M-S) rows contain only zeros.

The first problem to which we address ourselves is the construction of a vector K satisfying the inconsistency condition of Theorem 1. Let the 1x(R-T) vector $K_3' = (k_{T+1}, \ldots, k_R) = 0'$. Let the 1x(T-M) vector $K_2' = (k_{M+1}, \ldots, k_T)$ be an arbitrary positive vector. Let the 1x(M+1) vector $K_1' = (k_0, k_1, \ldots, k_M)$ be determined by $K_1' = -K_2' (B_2 - A_2 A_1^{-1} B_1 \vdots A_2 A_1^{-1})$. We note that K_1' is a non-negative vector for,

$$(K_2 > 0) \text{ and } (B_2' - A_2 A_1^{-1} B_1 \leq 0) \text{ and } (A_2 A_1^{-1} \lneq 0) \text{ imply } (K_1 \geq 0)$$

Furthermore we have $D'K = 0$ for

$$D'K = \begin{pmatrix} 1 & D_1' & B_2' & B_3' \\ 0 & A_1' & A_2' & A_3' \end{pmatrix} \begin{pmatrix} K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} (1 \vdots B_1')\,K_1 + B_2'\,K_2 \\ (0 \vdots A_1')\,K_1 + A_2'\,K_2 \end{pmatrix} \text{ since } K_3 = 0$$

Examining the two subvectors of D'K (transposed for convenience) we have:

$$K_1' \begin{pmatrix} 1 \\ \cdots \\ B_1 \end{pmatrix} + K_2'\,B_2 = -K_2'\,(B_2 - A_2A_1^{-1}B_1 \vdots A_2A_1^{-1}) \begin{pmatrix} 1 \\ \cdots \\ B_1 \end{pmatrix} + K_2'B_2$$

$$= K_2'(-(B_2 - A_2A_1^{-1}B_1 + A_2A_1^{-1}B_1) + B_2)$$

$$= K_2'\,(-B_2 + B_2) = 0$$

$$K_1' \begin{pmatrix} 0' \\ \cdots \\ A_1 \end{pmatrix} + K_2'\,A_2 = -K_2'\,(B_2 - A_2A_1^{-1}B_1 \vdots A_2A_1^{-1}) \begin{pmatrix} 0' \\ \cdots \\ A_1 \end{pmatrix} + K_2'A_2$$

$$= K_2'\,(-(0 + A_2A_1^{-1}A_1) + A_2) = K_2'\,(-A_2 + A_2) = 0$$

For application of theorem II we note that we have a $K \geq 0$ such that $D'K = 0$ with at least $k_1, \ldots, k_S$, $k_{M+1}, \ldots, k_T$ positive. Since the last M-S rows of $(A_1')^{-1}\,A_2'$ are zero $k_{S+1}, \ldots, k_M$ will be zero. In this case the problem posed by Theorem II is the existence of a U such that:

$$A_1(U) = \ldots = A_S\,(U) = A_{M+1}\,(U) = \ldots = A_T(U) = 0$$

which is equivalent to the problem of the existence of the U satisfying

$$A^*U + B^* = \begin{pmatrix} A_1^* \\ A_2^* \end{pmatrix} U + \begin{pmatrix} B_1^* \\ B_2^* \end{pmatrix} = 0$$

where $(a_{r1},\dots,a_{rN})$ for $r=1,\dots,S$ and $r=M+1,\dots,T$ make up the rows of A_1^* and A_2^* respectively while B_1^* and B_2^* are the corresponding constant vectors. The system is consistent if $r(A^*) = r(A^*\ B^*)$ and inconsistent if $r(A^*) < r(A^*\ B^*)$ where r denotes the rank of a matrix. In the tableau we have a U satisfying $A_1^*U + B_1^* = 0$ (see 3 below) so that $r(A_1^*) = r(A_1^*\ B_1^*)$, and we have expressed $(a_{r1},\dots,a_{rN})$ for $r=M+1,\dots,T$ as linear combinations of $(a_{m1},\dots,a_{mN})$ for $m=1,\dots,S$ so that $r(A^*) = r(A_1^*)$. Therefore $A^*U + B^* = 0$ is consistent if and only if every row of $(A_2^*\ B_2^*)$ is expressible as a linear combination of the rows of $(A_1^*\ B_1^*)$. If some $y_{0M+r} \neq 0$ $(r=M+1,\dots,T)$ in the tableau the corresponding $(b_r, a_{r1},\dots,a_{rN})$ is not expressible as a linear combination of the rows of $(A_1^*\ B_1^*)$; it also requires the unit vector of k_0. Then $r(A^*) < r(A^*\ B^*)$ and $A^*U + B^* = 0$ is inconsistent. If $y_{0M+r} = 0$ for $r=M+1,\dots,T$ we have the required linear combination, $r(A^*) = r(A^*\ B^*)$, we have a U such that $A^*U + B^* = 0$, and the constraints $A_1(U) \geq 0,\dots, A_S(U) \geq 0$, $A_{M+1}(U) \geq 0,\dots,A_T(U) \geq 0$ taken together imply one or more equality constraints on U.

In the latter case we will want to return to the original formulation of the problem to ascertain if these previously unrecognized equality constraints are acceptable in the context of the problem at hand. If they are acceptable we may partition $AU + B \geq 0$ into a system of equality constraints, $A^*U + B^* = 0$,

and a system of inequality constraints, $A_{M+1}(U) \geq 0, \ldots,$ $A_{N}(U) \geq 0, \ldots, A_{P+1}(U) \geq 0, \ldots, A_{R}(U) \geq 0$, and may apply the procedure of section I.

The same arguments apply to systems containing both strict and ordinary inequalities with the additional proviso that if any one of $A_1(U), \ldots, A_M(U)$, $A_{N+1}(U), \ldots, A_P(U)$ is a strict inequality constraint the system is inconsistent.

We note that an indication of an unbounded solution to the linear programming problem may appear before all artificial variables are replaced; in fact when $r(A) < M$ in an inconsistent system it must. The indication will appear as a nonbasic k_r for which all $y_{mM+r} \leq 0$ and $y_{M+1\,M+r} < 0$. In particular we must have $y_{mM+r} = 0$ for all artificial basic variables because the relative magnitudes of W, 1 and d are such that if all these $y_{mM+r} \leq 0$ and one of them is negative, $y_{M+1\,M+r} > 0$ which contradicts the hypothesis that we have an indication of an unbounded solution to the linear programming problem. Given such an indication we can construct $K \geq 0$ satisfying $D'K = 0$ by setting $k_m = -y_{mM+r}\,k_r \geq 0$ for basic k_m's where k_r 0 is arbitrary. Thus by Theorem I $AU + B > 0$ is inconsistent and if $y_{0M+r} \neq 0$ $AU + B \geq 0$ is also inconsistent.

III B 3. No artificial vectors in the basis and optimality conditions satisfied: The achievement of an optimal solution to the linear programming problem is signaled by the condition

$y_{M+1\,r} \geq 0$ for $r = 0, M+1, \ldots, M+R$

($y_{M+1\,r}$ for $r = 1, \ldots, M$ are irrelevant here.)

In this case we are assured by Theorems I and II that the system $AU + B \underset{(=)}{>} 0$ is consistent, and we can exhibit a first extreme point solution.

To show the extreme point solution we again assume for ease of discussion that we arranged the columns of the initial tableau so that $k_0, k_1, \ldots, k_M$ are the basic variables in the final tableau. With this assumption the series of row operations carried out to obtain this tableau is equivalent to the following transformation on the initial tableau.

$$\begin{pmatrix} 1 & -B_1'(A_1')^{-1} & 0 \\ 0 & (A_1')^{-1} & 0 \\ 0 & -C_1'(A_1')^{-1} & 1 \end{pmatrix} \begin{pmatrix} -((A_{22}^o)^{-1}B_2^o)' & 0 & 1 & 0' & B_1' & B_2' \\ -((A_{22}^o)^{-1}A_{21}^o)' & 0 & 0 & I_M & A_1' & A_2' \\ * & 0 & 0 & 0' & C_1' & C_2' \end{pmatrix} =$$

$$= \begin{pmatrix} ((A_{22}^o)^{-1}B_2^o)' - B_1'(A_1')^{-1}((A_{22}^o)^{-1}A_{21}^o)' & 0 & 1 & -B_1'(A_1')^{-1} & 0' & B_2' - B_1'(A_1')^{-1}A_2' \\ -(A_1')^{-1}((A_{22}^o)^{-1}A_{21}^o)' & 0 & 0 & (A_1')^{-1} & I_M & (A_1')^{-1}A_2' \\ * & 0 & 0 & -C_1'(A_1')^{-1} & 0' & C_2' - C_1'(A_1')^{-1}A_2' \end{pmatrix}$$

where $C_1' = B_1' - W\,1_M'\,A_1' - d\,1_M'$ is $1 \times M$ and corresponds to $k_1, \ldots, k_M$

$C_2' = B_2' - W\,1_M'\,A_2' - d\,1_{R-M}'$ is $1 \times (R-M)$ and corresponds to $k_{M+1}, \ldots, k_R$

We note that for r = 1,...,R: $y_{M+1,\ M+r} = y_{0M+r} + d(\sum_{m=1}^{m=M} y_{m,M+r} - 1)$. Thus if d is sufficiently small

$$(y_{0M+r} < 0) \text{ implies } (y_{M+1\ M+r} < 0) \quad \text{thus } (y_{M+1\ M+r} \geq 0) \text{ implies}$$

$$(y_{0M+r} \geq 0)$$

That is the satisfaction of the optimality conditions guarantees that $y_{0M+r} \geq 0$ for r = 1,...,R.

We now will demonstrate that $(y_{01},\ldots,y_{0M})$, the first row of the columns for the artificial variables, provide us with an extreme point solution of $AU + B \geq 0$. First we subtract "surplus" variables from the inequality constraints to convert the system to a set of equations of the form:

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} U - \begin{pmatrix} I_M & 0 \\ 0 & I_{R-M} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = 0$$

where X_1 is the Mx1 vector of surplus variables associated with the first M constraints and X_2 is the (R-M)x1 vector of surplus variables associated with the remaining R-M constraints. For a solution to satisfy all inequality constraints we must have $X \geq 0$. To obtain a solution with the first M constraints holding as equalities we set $X_1 = 0$, and make the transformation:

$$\begin{pmatrix} A_1^{-1} & 0 \\ 0 & I_{N-M} \end{pmatrix} \left[\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} U - \begin{pmatrix} I_M & 0 \\ 0 & I_{N-M} \end{pmatrix} \begin{pmatrix} 0 \\ X_2 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \right] = 0$$

$$\begin{pmatrix} I_M \\ A_2 \end{pmatrix} U + \begin{pmatrix} A_1^{-1} & 0 \\ 0 & I_{N-M} \end{pmatrix} \begin{pmatrix} 0 \\ X_2 \end{pmatrix} + \begin{pmatrix} A_1^{-1} B_1 \\ B_2 \end{pmatrix} = 0$$

which can be broken into two sets of equations:

$$U + A_1^{-1} B_1 = 0 \quad \text{or} \quad U' = -B_1' (A_1')^{-1}$$

$$A_2 U - X_2 + B_2 = 0 \quad \text{or} \quad X_2' = B_2' + U' A_2' = B_2' - B_1' (A_1')^{-1} A_2'$$

Referring to the matrix summary of the last tableau, we note that $(y_{01}, \ldots, y_{0M}) = -B_1' (A_1')^{-1} = U'$, that $(y_{0,M+1}, \ldots, y_{0N}) = B_2' - B_1' (A_1')^{-1} A_2' = X_2'$, and that the satisfaction of the optimality conditions implies $X_2' \geqq 0$. In addition we recall from section I that given a solution U, we can compute $U_2^o = -((A_{22}^o)^{-1} B_2^o + (A_{22}^o)^{-1} A_{21}^o U)$. In our last tableau we have:

$$(y_{0,-P}, \ldots, y_{0,-1}) = -\left([(A_{22}^o)^{-1} B_2^o]' - B_1'(A_1')^{-1} [(A_{22}^o)^{-1} A_{21}^o]' \right)$$

$$= (U_2^o)' \quad \text{for} \quad U' = -B_1' (A_1')^{-1}$$

Thus the linear programming technique has produced an extreme point solution to the original system of mixed constraints.

Although we have assumed a special ordering of the columns of the tableau for ease of exposition, the ordering of the columns of the identity matrix in the last tableau is irrelevant to the order of the elements of U' and $(U_2^0)'$. The positions of the various u's in these row vectors is fixed by the ordering of the rows of the tableau.

III B 4. Optimality conditions satisfied with artificial vectors remaining in basis: The appearance of this situation is a sufficient condition for $AU + B \underset{(\geq)}{>} 0$ to be consistent and unbounded. To facilitate the demonstration of this assertion, we assume that the rows and columns of the initial tableau were arranged so that $l_1, \ldots, l_S$ $(S < M)$ are the artificial variables which have been replaced by the legitimate variables $k_1, \ldots, k_S$. We note that an interchange of rows in the initial tableau dictates a corresponding interchange of elements in U. With this assumption the series of row operations is equivalent to the following transformation on the body of the tableau.

$$\begin{pmatrix} 1 & -B_1'(A_{11}')^{-1} & 0' & 0 \\ 0 & (A_{11}')^{-1} & 0 & 0 \\ 0 & -A_{12}'(A_{11}')^{-1} & I_{M-S} & 0 \\ 0 & -C_1'(A_{11}')^{-1} & 0' & 1 \end{pmatrix} \begin{pmatrix} 1 & 0' & 0' & B_1' & B_2' \\ 0 & I_S & 0 & A_{11}' & A_{21}' \\ 0 & 0 & I_{M-S} & A_{12}' & A_{22}' \\ 0 & 0' & 0' & C_1' & C_2' \end{pmatrix} =$$

$$\begin{pmatrix} 1 & -B_1'\,(A_{11}')^{-1} & 0' & 0' & B_2' - B_1'\,(A_{11}')^{-1}\,A_{21}' \\ 0 & (A_{11}')^{-1} & 0 & I_S & (A_{11}')^{-1}\,A_{21}' \\ 0 & -A_{12}'\,(A_{11}')^{-1} & I_{M-S} & 0 & A_{22}' - A_{12}'\,(A_{11}')^{-1}\,A_{21}' \\ 0 & -C_1'\,(A_{11}')^{-1} & 0' & 0' & C_2' - C_1'\,(A_{11}')^{-1}\,A_{21}' \end{pmatrix}$$

where $C_2' - C_1'\,(A_{11}')^{-1}\,A_{21} \geq 0$ signals the satisfaction of the optimality conditions.

We first will show that optimality conditions can be maintained while replacing artificial variables by legitimate variables in the following way. Note that in this discussion the index S will vary as more artificial variables are replaced.

Assume that $y_{m,\,M+r} > 0$ for some $m = S+1,\ldots,M$ and some $r = S+1,\ldots,R$. Then we select an artificial variable l_{m^*} for which this holds and replace it with k_{r^*} selected by the criterion

$$\frac{y_{M+1\;M+r^*}}{y_{m^*\;M+r^*}} = \min_{r=S+1,\ldots,R}\left(\frac{y_{M+1\;M+r}}{y_{m^*\;M+r}}\;\middle|\;y_{m^*\;M+r} > 0\right)$$

In the transformation we will replace $y_{M+1,\;M+r} \geq 0$ with

$$y_{M+1\;M+r} - y_{m^*\;M+r}\;\frac{y_{M+1\;M+r^*}}{y_{m^*\;M+r^*}} \geq 0$$

so that the optimality conditions are maintained. We continue replacing artificial vectors until we reach a situation in

which $y_{m,\,M+r} \leq 0$ for all $m = B+1,\ldots,M$ and $r = B+1,\ldots,R$. This must occur because with sufficiently large W it is not possible to satisfy the optimality conditions with only one artificial variable as a basic variable, say l_M, for which $y_{M,\,M+r} > 0$ for any $r = B+1,\ldots,R$ because in this case:

$$y_{M+1\;M+r} = y_{0\;M+r} + d\left(\sum_{m=1}^{M-1} y_{m\,M+r} - 1\right) - W\,y_{M\;M+r} \tilde{=} -W\,y_{M\;M+r}$$

Therefore $y_{M\;M+r} > 0$ implies $y_{M+1\;M+r} < 0$

Assume now that $y_{m*,\,M+r} < 0$ for some $m = B+1,\ldots,M$ and some $r = B+1,\ldots,R$. Then we select an artificial variable l_{m*} for which this holds and replace it with k_{r*} selected by the criterion

$$\frac{y_{M+1\;M+r^*}}{y_{m^*\;M+r^*}} = \underset{r=B+1,\ldots,R}{\text{Max}}\left(\frac{y_{M+1\;M+r}}{y_{m^*\;M+r}}\;\middle|\; y_{m^*\;M+r} < 0\right)$$

In the transformation we will replace $y_{M+1,\;M+r} \geq 0$ with

$$y_{M+1\;M+r} - y_{m^*\;M+r}\,\frac{y_{M+1\;M+r^*}}{y_{m^*\;M+r^*}} \geq 0$$

so that the optimality conditions will be maintained. We continue replacing until we (a) replace all artificial variables or (b) all $y_{m\,M+r} = 0$ for $m = S+1,\ldots,M$ and $r = S+1,\ldots,R$.

With regard to (a) we have a tableau of the form given in section 3 above, and in particular we note that in making the transformation which replaced say, l_M by k_M to obtain this tableau we replaced $y_{M\,M+r} \leq 0$ by $y_{M\,M+r}/y_{M\,M+M} \geq 0$ for $r = M,\ldots,R$, i.e. the last row of $(A_1')^{-1} A_2'$ in the new tableau contains only non-negative elements. As noted in section 3 we have a solution to $AU + B \geq 0$, and we now will exhibit the nontrivial solution to $AU \geq 0$ required by Theorem IV to establish that the set defined by $AU + B \underset{(\geq)}{>} 0$ is unbounded.

Let $U^* = A_1^{-1} e_M$ where $e_M' = (0,\ldots, 0, 1)$ is $1\times M$. Since $r(A_1) = M$ every column of A_1^{-1} contains at least one nonzero element so that $U^* \neq 0$.

$$AU^* = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} (A_1^{-1} e_M) = \begin{pmatrix} e_M \\ A_2A_1^{-1}e_M \end{pmatrix}$$

Clearly $e_M \geq 0$. As we noted above the last row of $(A_1')^{-1} A_2'$ contains only non-negative elements so that the last column of $A_2A_1^{-1}$ contains only non-negative elements. Therefore $A_2A_1^{-1}e_M \geq 0$, and U^* is a nontrivial solution of $AU^* \geq 0$.

With regard to (b) we have a tableau of the form given in this section with $A'_{22} - A'_{12} (A'_{11})^{-1} A'_{21} = 0$. Furthermore we have

$$v_{T+1\,M+r} = v_{0\,M+r} + d\left(\sum_{n=1}^{n=T} v_{n\,M+r-1}\right) \quad r = S+1,\ldots,M$$

and with d sufficiently small the optimality conditions assure that $B'_2 - B'_1 (A'_{11})^{-1} A'_{21} \geq 0$.

In this case we again have a solution of $AU + B \geq 0$ to be found in the first row of the column corresponding to the artificial variables. If we take $U' = (-B'_1 (A'_{11})^{-1} \mid 0'_{M-S})$ we have the following expression for AU + B.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} -A_{11}^{-1} B_1 \\ 0_{M-S} \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} -B_1 \\ -A_{21} A_{11}^{-1} B_1 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0_S \\ B_2 - A_{21} A_{11}^{-1} B_1 \end{pmatrix}$$

So we have a solution with the first S constraints holding as equalities. It is not an extreme point solution because the set possesses no extreme points. An extreme point in M-space is fixed by M linearly independent vectors, but the fact that we have expressed $(a_{S+1\,1},\ldots,a_{S+1\,M}),\ldots,(a_{N1},\ldots,a_{NM})$ in the tableau as linear combinations of $(a_{11},\ldots,a_{1M}),\ldots,(a_{S1},\ldots,a_{SM})$ implies that $r(A) = S < M$. Therefore there are not available M linearly independent vectors to specify an extreme point.

In addition the fact that $r(A) = r(A_{11})$ means that any solution U to $(A_{11}A_{12})\,U = A_{11}U_1 + A_{12}U_2 = 0$ will satisfy $AU = 0$. Such a solution is given by an arbitrary U_2 and $U_1 = -A_{11}^{-1}A_{12}U_2$. Thus there is a nontrivial solution of the homogeneous system $AU \geq 0$, and by Theorem IV, the set defined by $AU + B \underset{(\geq)}{>} 0$ is unbounded.

III C. A Necessary and Sufficient Condition for Unboundedness: We reiterate that the conditions of B 4 above are sufficient but not necessary for unboundedness. Therefore we will present a method for determining the existence of a nontrivial solution to $AU \geq 0$ if it exists. Before discussing the method we note that if $AU > 0$ is consistent $AU + B \underset{(\geq)}{>} 0$ is consistent while the converse is not true and that $AU \geq 0$ is always consistent since $U = 0$ satisfies it.

Theorems I, II, and III hold for homogeneous as well as non-homogeneous systems and consequently we test the consistency of $AU > 0$ with the linear programming problem.

$$\text{Max } z = (1,0,\ldots,0)\,Z \text{ subject to } Z \geq 0 \text{ and } D'Z = 0 \text{ where}$$

$$D = \begin{pmatrix} 1 & 0' \\ 0 & A \end{pmatrix}$$

Given a tableau for the system $AU + B \underset{(\geq)}{>} 0$, in particular the one with which we determined the system to be consistent, we can obtain a tableau for the system $AU > 0$ by setting

$B = 0$. On the tableau itself this is accomplished by subtracting the first row of the tableau (except for the column of k_o) from the last row (the $y_{M+1\ r}$ row) and then subtract the first row from itself (except for the column of k_o). By comparison with the tableau of section B 3 above with $B = 0$, we see that this row operation yields a tableau which is equivalent to the indicated transformation on the body of an initial tableau for the system $AU > 0$.

$$\begin{pmatrix} 1 & 0' & 0 \\ 0 & (A_1')^{-1} & 0 \\ 0 & -C_1'(A_1')^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0' & 0' & 0' \\ 0 & I_M & A_1' & A_2' \\ 0 & 0' & C_1' & C_2' \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0' & 0' & 0' \\ 0 & (A_1')^{-1} & I_M & (A_1')^{-1} A_2' \\ 0 & -C_1'\,(A_1')^{-1} & 0' & C_2' - C_1'\,(A_1')^{-1} A_2' \end{pmatrix}$$

where

$C_1' = -(W 1_M' A_1' + d\, 1_M')$ is $1 \times M$ and corresponds to $k_1, \ldots, k_M$

$C_2' = -(W 1_M' A_2' + d\, 1_{R-M}')$ is $1 \times (R-M)$ and corresponds to $k_{M+1}, \ldots, k_R$

Before proceeding we observe that in practice we would not subtract the first row from itself; we would leave it at the head of the tableau, and while we would ignore it

when applying the criterion for the basic variable to be replaced at each iteration, we would transform the row at each iteration in the usual manner. In addition we would continue to transform the left hand section of the tableau which corresponds to the original equality constraints as before. The additional computations entailed are relatively small and as long as the first row of the tableau for the columns of $k_1, \ldots, k_n$ remain non-negative, this tableau provides us with an extreme point solution of $AU + B \underset{(\geq)}{>} 0$ which would otherwise have to be obtained by the methods of section V. However to avoid complicating the discussion in this section we will assume that the subtraction has been made.

If after carrying out the subtraction, the new $C_2' - C_1'(A_1')^{-1} A_2' \geq 0$, the optimality conditions are satisfied. By Theorems I and III, $AU > 0$ is consistent so that there is a nontrivial solution of $AU \geq 0$, and by Theorem IV the set satisfying $AU + B \underset{(\geq)}{>} 0$ is unbounded.

If the optimality conditions are not satisfied we apply the simplex technique as in section A until we obtain satisfaction of the optimality conditions or an indication of an unbounded solution. In the latter case we have that $AU > 0$ is inconsistent, but we still have the possibility that $AU \geq 0$ admits to a nontrivial solution in which some constraints

hold as equalities. To ease the discussion we assume that foresight or good fortune permitted us to arrange the columns of the original tableau so that $k_0, k_1, \ldots, k_M$ are now the basic variables, so that $k_{M+1}, \ldots, k_T$ $(T \leq N)$ are the variables for which we have an indication of an unbounded solution, and more particularly so that $k_1, \ldots, k_S$ $(S \leq M)$ are the basic variables for which $y_{m,\, M+r} < 0$ for some $r = M+1, \ldots, T$. With this assumption the series of row operations carried out to yield the present tableau is equivalent to the following transformation on the body of the tableau.

$$\begin{pmatrix} 1 & 0' & 0 \\ 0 & (A_1')^{-1} & 0 \\ 0 & -C_1'(A_1')^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0' & 0' & 0' & 0' \\ 0 & I_M & A_1' & A_2' & A_3' \\ 0 & 0 & C_1' & C_2' & C_3' \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0' & 0' & 0' & 0' \\ 0 & (A_1')^{-1} & I_M & (A_1')^{-1}A_2' & (A_1')^{-1}A_3' \\ 0 & -C_1'(A_1')^{-1} & 0' & C_2'-C_1'(A_1')^{-1}A_2' & C_3'-C_1'(A_1')^{-1}A_3' \end{pmatrix}$$

where

$C_1' = -(W\, 1_M'\, A_1' + d\, 1_M')$ is $1 \times M$ and corresponds to $k_1, \ldots, k_M$

$C_2' = -(W\, 1_M'\, A_2' + d\, 1_{T-M}')$ is $1 \times (T-M)$ and corresponds to $k_{M+1}, \ldots, k_T$

$C_3' = -(W\, 1_M'\, A_3' + d\, 1_{N-T}')$ is $1 \times (N-T)$ and corresponds to $k_{T+1}, \ldots, k_N$

and where for an indication of an unbounded solution we have

$$C_2' - C_1' (A_1')^{-1} A_2' < 0$$

$$(A_1')^{-1} A_2' \leq 0$$ where the last M-S rows contain only zeros.

We construct a $K \geq 0$ satisfying $D'K = 0$ by setting $K_3' = (k_{T+1},\ldots,k_R) = 0'$, by letting $K_2' = (k_{M+1},\ldots,k_T)$ be an arbitrary positive vector, and by setting $K_1' = (k_0,\ldots,k_M) = -K_2' (0 \mathrel{\vdots} A_2 A_1^{-1})$. As was shown in section E 2 this provides a $K \geq 0$ satisfying $D'K = 0$ with $k_1,\ldots,k_S$, $k_{M+1},\ldots,k_T$ positive. In this case the problem posed by Theorem II is the existence of a nontrivial U satisfying:

$$A_1^* (U) = \ldots = A_S^* (U) = A_{M+1}^* (U) = \ldots = A_T^* (U) = 0$$

where

$$A_r^* (U) = \sum_{m=1}^{m=M} a_{rm} u_m$$

which is equivalent to the problem of the existence of a nontrivial U satisfying:

$$A^* U = \left(\frac{A_1^*}{A_2^*} \right) U = 0$$

where the rows of A_1^* are $(a_{11},\ldots,a_{1M}),\ldots,(a_{S1},\ldots,a_{SM})$ in that order, and the rows of A_2^* are $(a_{M+1\,1},\ldots,a_{M+1\,M}),\ldots,(a_{T1},\ldots,a_{TM})$ in that order.

Since the last (N-R) rows of $(A_1')^{-1} A_2'$ in the tableau are zero, we have expressed in the tableau each row of A_2^* as a linear combination of the rows of A_1^*. Therefore $r(A^*) = r(A_1^*) = R$ and any U satisfying $A_1^* U = 0$ will also satisfy $A^* U = 0$. From the general theory of simultaneous linear equations in N unknowns, we have that the $A^* U = 0$ has a unique solution (in this case $U = 0$) if $r(A^*) = N$ and has an infinite number of solutions (consequently non-trivial ones) if $r(A_1^*) < N$. Therefore if $R < N$ or equivalently if there exists a basic variable k_m for which $y_{m\,N+r} = 0$ for $r = N+1,\dots,T$, then $AU \geqq 0$ admits to a non-trivial solution and the set satisfying $AU - B \overset{>}{(\geqq)} 0$ is unbounded.

In summary the application of the procedure of this section will provide necessary and sufficient conditions for the convex set to be unbounded. It will be pointed out in section V that even if this procedure is not applied the process of generating extreme points will eventually reveal unboundedness. Therefore an alternative to this procedure is to rely on the results of sections III B 4 and V.

IV. Testing Constraints for Redundancy: A constraint is redundant if every point satisfying the remaining constraints must satisfy the one in question. The identification of

redundant constraints may be an end in itself or an adjunct to the generation of the extreme points of a convex set. In the latter case the presence of redundant constraints is undesirable because it means that we will be carrying out calculations for the columns of the tableau corresponding to the redundant constraints when in fact they are useless for the determination of extreme points. Therefore it is desirable to eliminate redundant constraints from the problem if this can be done in such a way as to produce a net savings in computational effort. The method below involves the application of Theorems VI, VII and VIII, and usually will produce a savings because the computations required to check redundancy are ones which would be required anyway in the course of generating extreme points.

IV A. General Method for the Redundancy Test: By the time we have reached the stage of applying this method we will have established that $AU + B > 0$ is consistent, that $r(A) = M$, and for nonbasic variables, say $k_{M+1},\ldots,k_R$, that $r(A)$ is not decreased by the removal of the r-th row of A for $r = M+1,\ldots,R$. Therefore we are in a position to test one of $A_{M+1}(U),\ldots,A_R(U)$ for redundancy; for notational convenience assume that we test $A_R(U)$.

From the result of section III B 3 we have an "optimal" tableau the body of which has the form shown on the right.

$$\begin{pmatrix} 1 & -B_1'\,(A_1')^{-1} & 0 \\ 0 & (A_1')^{-1} & 0 \\ 0 & -C_1'\,(A_1')^{-1} & 1 \end{pmatrix}\begin{pmatrix} 1 & 0' & B_1' & B_2' & B_3' \\ 0 & I_M & A_1' & A_2' & A_3' \\ 0 & 0' & C_1' & C_2' & C_3' \end{pmatrix} =$$

$$\begin{pmatrix} 1 & -B_1'(A_1')^{-1} & 0' & B_2'-B_1'(A_1')^{-1}A_2' & B_3'-B_1'(A_1')^{-1}A_3' \\ 0 & (A_1')^{-1} & I_M & (A_1')^{-1}A_2' & (A_1')^{-1}A_3' \\ 0 & -C_1'(A_1')^{-1} & 0' & C_2'-C_1'(A_1')^{-1}A_2' & C_3'-C_1'(A_1')^{-1}A_3' \end{pmatrix}$$

where

$C_1' = (B_1' - W\,1_M'\,A_1' - d\,1_M')$ is 1 x M and corresponds to $k_1, \ldots, k_M$

$C_2' = (B_2' - W\,1_M'\,A_2' - d\,1_{R-M-1}')$ is 1 x (R-M-1) and corresponds to $k_{M+1}, \ldots, k_{R-1}$

$$C_3' = (B_3' - W\,1_M'\,A_3' - d) = b_R - W\sum_{m=1}^{m=M} a_{Rm} - d$$

corresponds to k_R.

and $(C_2' - C_1'(A_1')^{-1}A_2' \; \vdots \; C_3' - C_1'\,(A_1')^{-1}A_3') \geq 0$ by the optimality conditions.

In order to seek the conditions of Theorem VII for $A_R(U)$, we wish to establish if the following linear programming problem has a basic feasible solution.

$$\text{Max } z = (1,d,\ldots,d)\,K \text{ subject to } K \geq 0\,,\; B'K = \begin{pmatrix} b_R \\ a_{R1} \\ \vdots \\ a_{RM} \end{pmatrix}$$

where now,

$$K' = (k_0, k_1, \ldots, k_{R-1})$$

$$D' = \begin{pmatrix} 1 & B_1' & B_2' \\ 0 & A_1' & A_2' \end{pmatrix}$$

This is the linear programming problem we would have obtained from our original problem had we removed the last column of the initial tableau and inserted it in place of the zero requirement vector.

In order to start to determine if this linear programming problem has a feasible solution, we note that had we applied the same series of row operations to a tableau for this problem as we applied to the tableau of the original problem, we would have arrived at a tableau which is the one given above; the only difference would be that now the last column of the tableau is the transformed requirement vector which in a linear programming problem gives the values of the basic variables in the solution for the new linear programming problem. It is an optimal solution since $C_2' - C_1' (A_1')^{-1} A_2' \geq 0$, and if $y_{m\ N+R} \geq 0$ for $m = 0,1,\ldots,M$ it is also feasible giving the conditions which by Theorem VII imply that $A_R(U)$ is redundant. If some $y_{m\ N+R} < 0$ the solution is not feasible, and we must use some device to determine if a basic feasible optimal solution exists. The dual simplex algorithm is a suitable device. For a full

III A. We note that these criteria will insure that the optimality conditions are maintained after the transformation. We also note that the first row of the columns of the tableau corresponding to the artificial variables will give us U of the extreme point solution for which $A_1(U) = \ldots = A_{m^*-1}(U) = A_{r^*}(U) = A_{m^*+1}(U) = \ldots = A_M(U) = 0$. If we also carry out the row operations on the left hand section of the tableau we obtain the corresponding U_2^0 .

The repeated application of this procedure either will provide a basic feasible optimal solution in which case we have that $A_R(U)$ is redundant or will lead to a situation where we have a k_{m^*} for which $y_{m^* N+R} < 0$ but $y_{m^* N+r} \geq 0$ for $r = N+1, \ldots, N-1$ in which case the linear programming problem has no basic feasible solution and $A_R(U)$ is not redundant.

In general a linear programming problem has no feasible solution, a basic feasible optimal solution, or an unbounded solution. We have considered the first two possibilities above. An unbounded solution is not possible in this case because we started with a tableau for which $y_{M+1\,N+r} = z_r - c_r \geq 0$ for $r = 1, \ldots, N-1$ (Hadley (1962) p. 347).

IV C. Other Criteria for Redundancy and Nonredundancy of Constraints: The procedure above can be used to check any constraint $A_s(U)$ in the consistent system $AU + B \overset{>}{(\geq)} 0$ where $r(A) = N$ if the removal of $(a_{s1}, \ldots, a_{sN})$ does not

decrease the rank of A. For an $A_s(U)$ such that the removal of $(a_{s1},\dots,a_{sN})$ will decrease the rank of A, 1. below will eventually establish that $A_s(U)$ is not redundant. The other criteria are included as optional criteria to shorten the process of checking for redundancy.

1. Application of Theorem VIII: If in any tableau in which $k_0, k_1,\dots,k_M$ are basic variables, $y_{0M+r} > 0$ for $r = M+1,\dots,N$, then by Theorem VIII none of $A_1(U),\dots,A_M(U)$ is redundant and need not be checked by the procedure above. Recall that $y_{0M+r} > 0$ implies that at the extreme point in question $A_r(U) > 0$. In particular if $r(A)$ will be decreased by the removal of any, $(a_{i1},\dots,a_{iN})$, k_i will never leave the basis using the procedure above and eventually when all redundant constraints are discovered we will have this condition satisfied for $A_i(U)$.

2. Identical constraints: If in any tableau we have two columns which are identical, then one of them may be deleted since the two constraints are in reality the same.

3. Application of Theorem VI: Clearly it is not necessary to go through the formality of applying the dual simplex algorithm if an equivalent result is more readily obtainable. If after the removal of all artificial vectors from the basis we have in the tableau a nonbasic k_r for which $y_{m\,M+r} \geq 0$ for $m = 0,1,\dots,M$ we have produced the set of constants which by

Theorem VI is sufficient to declare that $A_r(U)$ is redundant.

We assume that this procedure is used to eliminate all redundant constraints, that a suitable condensation of the tableau has been made, and that we have renumbered so that we have a convex set defined by R nonredundant constraints.

V. Generating Extreme Points: By the time we have reached the stage of generating the extreme points of the set satisfying $AU + B \underset{(\geq)}{>} 0$, we will have established that $AU + B > 0$ is consistent, that $r(A) = M$, that none of the remaining constraints are redundant, and if we used the procedure of section III C whether or not the set is bounded.

V A. Recording Extreme Points: A reasonably compact device for recording relevant information about each extreme point is a table with R columns which correspond to the constraints (the constraints appear in the table in the same order as in the tableau) and $\binom{R}{M}$ rows corresponding to all combinations of the R constraints taken M at a time. The number of extreme points will usually be considerably less than this number. A particular row will have R-M of its cells blocked out; the M unblocked cells correspond to the particular constraints which hold as equalities at that point. We may arrange the rows so that the first $\binom{R-1}{M-1}$ rows are for points at which the first constraint holds as an equality while the

last $\binom{R-1}{M}$ rows are for points at which the first constraint holds as a strict inequality. Within the first set determined above we arrange the rows so that in the first $\binom{R-2}{M-2}$ both the first and second constraints hold as equalities while in the second $\binom{R-2}{M-1}$ rows the second constraint holds as a strict inequality. In the set of rows for which the first constraint holds as a strict inequality, we arrange the rows so that in the first $\binom{R-2}{M-1}$ the second constraint holds as an equality and in the second $\binom{R-2}{M}$ rows the second constraint holds as a strict inequality. We repeat the same process for each of the four sets determined above for the third constraint, etc. until we have identified all $\binom{R}{M}$ points by the constraints which hold as equalities. If we had P equality constraints in the original problem and chose to carry a tableau with a left hand section for these constraints, we add P more columns to the table for these.

An example of such a table is given below for M = 3, P = 2, and R = 5.

Constraints						
1	2	3	4	5	1	2
			X	X		
		X		X		
		X	X			
	X			X		
	X		X			
	X	X				
X				X		
X			X			
X		X				
X	X					

If in the course of our calculations we find that a row corresponds to an extreme point, we enter $u_1, \ldots, u_M$ for that extreme point from left to right in the open cells of that row and the elements of the corresponding U_2^o in the P columns of that row on the right of the table if appropriate. When we find that the point represented by a row lies outside the convex set, we strike that row from the table.

Of course one may start such a table before checking for redundancy so that the extreme points found in the course of the redundancy check may be entered immediately. In this case when a constraint is found to be redundant the corresponding column of the table and any rows having unblocked cells in that column are struck from the table.

V B. Determining Adjacent Extreme Points: We now turn to the problem of determining the extreme points which are "adjacent" to the extreme point given by a particular tableau in the sense that they share an edge of the convex set with the given extreme point. In general if we have an extreme point at which $V \geq M$ constraints hold as equalities there may be no adjacent extreme point or as many as $\binom{V}{M-1}$ adjacent extreme points, the precise number of adjacent extreme points depending on the geometry of the convex set defined by $AX + B \underset{(=)}{>} 0$.

For ease of discussion we assume that in the tableau at hand $k_0, k_1, \ldots, k_M$ are the basic variables and that

$$A_1(U) = \ldots = A_V(U) = 0 \qquad (V \geq M)$$

at the extreme point given by this tableau, i.e. $y_{0M+r} = 0$ for $r = 1, \ldots, V$. For each k_m $(m=1, \ldots, M)$ we may compute

$$w_{mt} = y_{0M+t_m} / y_{mM+t_m} = \max_{V+1 \leq r \leq R} \left(y_{0M+r} / y_{mM+r} \mid y_{mM+r} < 0 \right)$$

and for each m there are four possibilities.

1. $w_{mt} < 0$ is unique. Then we may generate an adjacent extreme point by replacing k_m with k_{t_m} using the transformation procedure of section III 4. The new extreme point is given by $(y_{01}, \ldots, y_{0M})$ in the new tableau, and if k_{t_m} is chosen by the given criterion this point will satisfy all constraints at least as ordinary inequalities because in the transformation we will replace y_{0M+r} by

$$y_{0M+r} - y_{mM+r} \, (w_{mt}) \qquad \text{for } r = 1, \ldots, R$$

Now $w_{mt} < 0$, and $y_{0M+r} \geq 0$ for $r = 1, \ldots, R$ because we apply the criterion in a tableau corresponding to an extreme point. If $y_{mM+r} < 0$ $(t_m \neq r = 1, \ldots, R)$ we have the following inequality because w_{mt} is unique

$$w_{mt} > (y_{0N+r}/y_{mN+r}) \qquad r \neq t_m$$

$$y_{mN+r}w_{mt} < y_{0N+r}$$

$$0 < y_{0N+r} - y_{mN+r}w_{mt}$$

If $y_{mN+r} \leq 0$ ($t_m \neq r = 1,\ldots,R$) we have immediately that

$$y_{0N+r} - y_{mN+r}w_{mt} \geq y_{0N+r} \geq 0$$

Therefore all constraints, which for purposes of specifying a point on the boundary of the set must be regarded as ordinary inequalities, will be satisfied after the transformation.

2. $w_{mt} < 0$ is not unique. In this case any one of the candidates for the maximum may be selected as k_{t_m}. For the remaining k_r's tied for the maximum we have

$$y_{0N+r} - y_{mN+r}w_{mt} = 0$$

Thus all such constraints will hold as equalities at the extreme point found when k_{t_m} replaces k_m as a basic variable.

3. The criterion cannot be applied because $y_{mN+r} \leq 0$ for $r = N+1,\ldots,R$. In section III B 4 we demonstrated that the existence of a basic variable, which we called k_s there, having $y_{mN+r} \leq 0$ for $r = N+1,\ldots,R$ is a sufficient condition for declaring the set defined by $AU + B \underset{(\geq)}{>} 0$ to be unbounded because we can exhibit a nontrivial solution to $AU \geq 0$. Here we will examine the nature of this unboundedness more closely.

first noting that the N-1 equations

$$A_1(U) = \dots = A_{m-1}(U) = A_{m+1}(U) = \dots = A_N(U) = 0$$

fix a line, L(U), in N-space and that the intersection of L(U) and $A_m(U) = 0$ fixes an extreme point of the set.

We now set

$$U = -(A_1^{-1}B_1) + g(A_1^{-1}e_m)$$

where $g > 0$ and $e_m' = (0,\dots,0,1,0,\dots,0)$ is a 1xN unit vector having the one in its m-th position. That is we let U be the vector sum of the extreme point at hand and a scalar multiple of one nontrivial solution to $AU \geq 0$ (cf. section III B 4) so that

$$A_1U + B_1 = A_1(-A_1^{-1}B_1) + gA_1A_1^{-1}e_m + B_1 = B_1 - B_1 + ge_m = ge_m \geq 0$$

Therefore this U is a half-line, $L^*(U)$, which is the half of L(U) satisfying $A_m(U) > 0$. Furthermore this half-line is an edge of the convex set because it satisfies the remaining M-N constraints which for purposes of defining a part of the boundary of the set need be regarded as ordinary inequalities. The latter statement is verified by examining

$$A_2U+B_2 = A_2(-A_1^{-1}B_1)+gA_2A_1^{-1}e_m+B_2 = B_2-A_2A_1^{-1}B_1+gA_2A_1^{-1}e_m$$

We have $B_2 - A_2A_1^{-1}B_1 \geq 0$ in the tableau at hand and $A_2A_1^{-1}e_m \geq 0$ by virtue of the condition that $y_{m,N+r} \geq 0$ for

r = M+1,...,N, i.e. the r-th column of $A_2A_1^{-1}$ contains no negative entries.

If the convex set is unbounded, recalling that r(A) = N, there must be at least one extreme point which lies on a line such as L(U), and at least one extreme point on each edge of the convex set. Therefore a procedure that attempts to determine the extreme points which are adjacent to every extreme point of the set will eventually uncover the fact that a convex set is unbounded and if continued long enough will produce every edge of the form $L^*(U)$. When r(A) = N the unbounded convex set defined by $AU + B \underset{(\geq)}{>} 0$ can also be defined by a list of its extreme points and its half-line edges because every point in the set can be expressed as a convex combination of the extreme points and points on each of the half-line edges such as $L^*(U)$.

4. $w_{mt} = 0$. This situation must arise in any tableau corresponding to an extreme point where $V \geq N+1$ constraints hold as equalities. For the sake of definiteness let us say that $A_V(U) = 0$ at the point in question where $V \geq N+1$. Then $y_{0,M+V} = 0$ in our tableau, and for any k_m such that $y_{m,M+V} < 0$ we will have $w_{mt} = 0$. Such a k_m must exist because if $y_{m,M+V} \geq 0$ for m = 1,...,N $A_V(U) \underset{(\geq)}{>} 0$ would have been discarded as being redundant by Theorem VI. Since $w_{mt} = 0$ the replacement of k_m by k_{t_m} leaves us at the same

extreme point although we will have a different tableau. In general $\binom{V}{M}$ bases (and thus $\binom{V}{M}$ tableaus) can be used to represent each extreme point. When $V \geq M+1$, $\binom{V}{M} > 1$ and we must consider the possibility that there are edges emanating from the extreme point at hand which cannot be detected by replacing one of the M basic variables in a particular tableau which we have used to represent the point. Because each edge of the convex set emanating from the given extreme point is specified by M-1 of the V equations which hold at the point together with one or two inequalities according as the set is unbounded or bounded, it is sufficient (although not usually necessary) to apply the procedure of VC below to each of the tableaus which can represent the extreme point in order to detect all such edges and therefore to determine all adjacent extreme points.

V C. Procedure: We start with a table in which all extreme point solutions found in the course of applying the procedures of sections III and IV have been entered. We take one of the previously computed tableaus which corresponds to an extreme point. Noting that the last row of the tableau ($y_{M+1,r}$'s) and the "value" information (z and d's) are not needed for this procedure, we discard them.

1. For each basic variable in the tableau (say k_n where $n = 1,\ldots,M$) we compute w_{nt}.

If $u_{mt} < 0$ is unique we strike from the table those rows in which the M unblocked cells are in columns corresponding to $k_1, \dots, k_{m-1}, k_{m+1}, \dots, k_M$ and k_r ($M+1 \leq r \neq t_m$). We do this because the M-1 equations

$$A_1(U) = \dots = A_{m-1}(U) = A_{m+1}(U) = \dots = A_M(U) = 0$$

fix a line, L(U), on which there are at most two extreme points. The intersections of L(U) with $A_m(U)=0$ and $A_{t_m}(U)=0$ give these two points. Therefore the intersections of L(U) with $A_r(U)=0$ ($M+1 \leq r \neq t_m$) must specify exterior points.

If $u_{mt} < 0$ does not uniquely determine a $t_m \geq M+1$, we may select as k_{t_m} any of the k_r's which are tied for the maximum and we strike from the table those rows in which the unblocked cells are in columns corresponding to $k_1, \dots, k_{m-1}$, $k_{m+1}, \dots, k_M$ and k_r ($r \geq M+1$) where k_r is not one of those tied for the maximum because these rows correspond to exterior points.

In either case if the extreme point at which

$$A_1(U) = \dots = A_{m-1}(U) = A_{m+1}(U) = \dots = A_M(U) = A_{t_m}(U) = 0$$

has not been entered in the table, we transform the tableau using the procedure of section III A to obtain the point.

If $u_{mt} = 0$ the transformation is carried out and yields a new representation of the same extreme point.

When w_{mt} cannot be computed because there is no $r \geq M+1$ for which $y_{m,M+r} < 0$, we strike from the table rows in which the M unblocked cells are in columns corresponding to $k_1,\dots,k_{m-1},k_{m+1},\dots,k_M$ and k_r ($r \geq M+1$) and make a note that the M-1 equations

$$A_1(U) = \dots = A_{m-1}(U) = A_{m+1}(U) = \dots = A_M(U) = 0$$

and the inequality $A_m(U) \underset{(\geq)}{>} 0$ define a half-line edge of the convex set on which there is only one extreme point.

2. When we first obtain a tableau such that $y_{0,M+r} = 0$ for one or more $r = M+1,\dots,N$ (say $r = M+1,\dots,V$) we strike from the table the M(V-M) rows in which the unblocked cells are in columns corresponding to k_r ($r = M+1,\dots,V$) and $k_1,\dots,k_M$ less k_m ($m = 1,\dots,M$) and make a note in the row in which the M unblocked cells are in the columns corresponding to $k_1,\dots,k_M$ that the constraints corresponding to $k_{M+1},\dots,k_V$ also hold as equalities at this point. Then no other entries are made in the table as other tableaus are produced for studying that particular extreme point.

3. We repeat the procedure for every tableau including those which are generated in 1. above until every adjacent extreme point has been computed for every tableau. When this is accomplished we will have all extreme point solutions entered in the table, and every row of the table must contain a solution or must be lined through.

VI. Incorporation of New Constraints: We suppose that at some stage of the procedure we wish to restrict U more stringently than was done by the original system of constraints by adding the constraints

$$A_{N+v}(U) \underset{(=)}{>} 0 \qquad v = 1,\ldots,V$$

to the original system. These V constraints can be expressed as

$$A_3 U + B_3 \underset{(=)}{>} 0$$

where A_3 is VxM and B_3 is Vx1.

In order to see how the new constraints can be incorporated into the system, we note that had we anticipated this situation we could have incorporated these constraints at the beginning by formulating the body of our initial tableau as

$$T = \begin{pmatrix} 1 & 0' & 0' & B_1' & B_2' & B_3' \\ 0 & I_S & 0 & A_{11}' & A_{21}' & A_{31}' \\ 0 & 0 & I_{N-S} & A_{12}' & A_{22}' & A_{32}' \\ 0 & 0' & 0' & C_1' & C_2' & C_3' \end{pmatrix}$$

Had we done this and applied the simplex algorithm subject to the stipulation that $k_{N+1},\ldots,k_{N+V}$ could not become basic variables, we would have carried out the same series of row operations as we actually did in order to progress to the stage at which we are now attempting to incorporate the new

constraints. With a convenient arrangement of rows and columns in our initial tableau this series of row operations is equivalent to the transformation on the body of the tableau.

$$\begin{pmatrix} 1 & -B_1'(A_{11}')^{-1} & 0' & 0 \\ 0 & (A_{11}')^{-1} & 0 & 0 \\ 0 & -A_{12}'(A_{11}')^{-1} & I_{m-r} & 0 \\ 0 & -C_1'(A_{11}')^{-1} & 0' & 1 \end{pmatrix} T =$$

$$\begin{pmatrix} 1 & -B_1'(A_{11}')^{-1} & 0' & 0' & B_2'-B_1'(A_{11}')^{-1}A_{21}' & B_3'-B_1'(A_{11}')^{-1}A_{31}' \\ 0 & (A_{11}')^{-1} & 0 & I_r & (A_{11}')^{-1}A_{21}' & (A_{11}')^{-1}A_{31}' \\ 0 & -A_{12}'(A_{11}')^{-1} & I_{m-r} & 0 & A_{22}'-A_{12}'(A_{11}')^{-1}A_{21}' & A_{32}'-A_{12}'(A_{11}')^{-1}A_{31}' \\ 0 & -C_1'(A_{11}')^{-1} & 0' & 0' & C_2'-C_1'(A_{11}')^{-1}A_{21}' & C_3'-C_1'(A_{11}')^{-1}A_{31}' \end{pmatrix}$$

Our illustrative transformation assumes that some artificial variables remain as basic variables when we wish to add constraints. An analogous example holds if all artificial vectors have been removed before we add new constraints.

This matrix transformation suggests the following procedure for incorporating additional constraints in the tableau at any stage.

1. For the constraints to be added form the matrix $(B_3 \vdots A_3)'$ insuring that the a corresponding to the n-th

column of A_3 is the same as that which corresponds to the n-th column of A_1 and A_2.

2. Consider the first M+1 rows (i.e. all except the last) of the M+1 columns in the tableau where the identity matrix was located in the initial tableau as an (M+1)x(M+1) matrix G'. Thus in our example above we would take

$$G' = \begin{pmatrix} 1 & -B_1'\,(A_{11}')^{-1} & 0' \\ 0 & (A_{11}')^{-1} & 0 \\ 0 & -A_{12}'\,(A_{11}')^{-1} & I_{M-N} \end{pmatrix}$$

In general G' will not be so conveniently arranged. However the arrangement is irrelevant; we simply use G' as we find it in the tableau.

3. To obtain the columns of $(B_3 \vdots A_3)'$ as linear combinations of the basic vectors (the form in which they must be incorporated) we perform the matrix multiplication

$$G'\,(B_3 \vdots A_3)'$$

4. We obtain the (M+1)-st row of the last V columns of the new tableau by the standard linear programming calculation

$$y_{M+1\;M+N+v} = \sum_{m=0}^{M+1} y_{m,-(M+1)}\, y_{m\;M+N+v} - d$$

for $v = 1,\ldots,V$

Note that we assign the "value" 0 to each of $x_{N+1},\ldots,x_{N+V}$.

Once the new constraints are incorporated in the tableau we can proceed as if they had been present from the beginning. If some $y_{m+i,\,n+m+v} < 0$, k_{n+v} is a candidate for basic variable at the next iteration. We note that even if the original system has been found to be consistent, the augmented system may be inconsistent. If the set defined by the original system was bounded so must be the set defined by the new system although the converse is not true. Some of the newly added constraints may be redundant or their inclusion may render a constraint in the original system redundant.

VII. Example: To illustrate the method we will apply it to data reported by Gillison (1954). This article describes the utilization of a variant of the Churchman-Ackoff technique for weighting qualitative objectives (Chapter 2) for the specification of the relative importance of six types of defects which could appear in lots of a pharmaceutical product. A board composed of nine men was presented with a series of judgments of the sort, "Is a type 'a' defect more or less important than type 'b' and 'c' defects taken together?" In each case a vote was taken, the decision of the majority was accepted, and adjustments were made in the tentative values assigned to the defects to bring them into accord with all the comparisons made. Letting v_a through v_f represent the relative importance of the six types of

defects, we can condense the results given by Stillson (1954) into the twelve relations below.

1. $v_a > v_b$
2. $v_b > v_c$
3. $v_c > v_d$
4. $v_e > v_f$
5. $v_f > 0$
6. $v_a < v_b + v_e$
7. $v_a > v_b + v_f$
8. $v_b < v_c + v_d + v_e$
9. $v_b > v_c + v_e$
10. $v_c < v_d + v_f$
11. $v_b = v_c + v_d$
12. $v_d = v_c + v_f$

We anticipate the results to be obtained later by noting here that this list can be condensed further. The decision was made to set $v_a = 100$ and to express the other values as percentages of v_a.

VII A. Procedure of Section I: We use the decision $v_a = 100$ and the two equalities to reduce the number of variables which need to be considered explicitly in generating extreme points. Following section I we set $U^{o'} = (U_1^{o'} \;\; U_2^{o'}) = (v_f v_e v_c \;\; v_d v_b)$ and we have:

$$\begin{pmatrix} A_1^o & B_1^o \\ A_2^o & B_2^o \end{pmatrix} = \left|\begin{array}{rrrrr|r|l} 0 & 0 & 0 & 0 & -1 & 100 & \qquad (1) \\ 0 & 0 & -1 & 0 & 1 & 0 & \qquad (2) \\ 0 & 0 & 1 & -1 & 0 & 0 & \qquad (3) \\ -1 & 1 & 0 & 0 & 0 & 0 & \qquad (4) \\ 1 & 0 & 0 & 0 & 0 & 0 & \qquad (5) \\ 0 & 1 & 0 & 0 & 1 & -100 & \qquad (6) \\ -1 & 0 & 0 & 0 & -1 & 100 & \qquad (7) \\ 0 & 1 & 1 & 1 & -1 & 0 & \qquad (8) \\ 0 & -1 & -1 & 0 & 1 & 0 & \qquad (9) \\ 1 & 0 & -1 & 1 & 0 & 0 & \qquad (10) \\ \hline -1 & -1 & 0 & 1 & 0 & 0 & \qquad (11) \\ 0 & 0 & -1 & -1 & 1 & 0 & \qquad (12) \end{array}\right.$$

The matrix given below is the matrix

$$\begin{pmatrix} A = A^{o}_{11} - A^{o}_{12}\,(A^{o}_{22})^{-1}A^{o}_{21} & 0 & B = B^{o}_{1} - A^{o}_{12}\,(A^{o}_{22})^{-1}B^{o}_{2} \\ (A^{o}_{22})^{-1}\,A^{o}_{21} & I_2 & (A^{o}_{22})^{-1}\,B^{o}_{2} \end{pmatrix}$$

and is obtained by performing row operations on the rows of the first matrix as indicated to the right of each row.

$$\left|\begin{array}{ccc|cc|c} -1 & -1 & -1 & 0 & 0 & 100 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & -100 \\ -2 & -1 & -1 & 0 & 0 & 100 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & -1 & 0 & 0 & 0 \\ \hline -1 & -1 & 0 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 1 & 0 \end{array}\right| \qquad \begin{array}{l} (1) = (1) + (11) + (12) \\ (2) = (2) - (11) - (12) \\ (3) = (3) + (11) \\ (4) = (4) \\ (5) = (5) \\ (6) = (6) - (11) - (12) \\ (7) = (7) + (11) + (12) \\ (8) = (8) + (12) \\ (9) = (9) - (11) - (12) \\ (10) = (10) - (11) \\ (11) = (11) \\ (12) = (12) + (11) \end{array}$$

We note that the incorporation of the equality constraints into the system of inequalities has illuminated some redundancies. For example (4) and (5) imply (2) and (8) while (5) and (9) are identical. Therefore we can discard constraints (2), (8) and (9), redefining A and B accordingly.

By judicious searching we might discover other redundancies, but we will rely instead on the formal procedure to reveal any other redundancies.

VII B. Procedure of Section III: We form our initial tableau from the last matrix above with k_1 corresponding to row 1, k_2 to row 3, k_3 to row 4, k_4 to row 5, k_5 to row 6, k_6 to row 7, and k_7 to row 10.

First Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_4	k_5	k_6	k_7
				1	$-M$	$-M$	$-M$	d	d	d	d	d	d	d
k_0	1	0	0	1	0	0	0	100	0	0	0	-100	100	0
l_1	$-M$	1	1	0	1	0	0	-1	-1	1	1	1	-2	2
l_2	$-M$	1	1	0	0	1	0	-1	-1	1	0	2	-1	1
l_3	$-M$	0	1	0	0	0	1	-1	1	0	0	1	-1	-1
				0	0	0	0	$3M + 100 - d$	$M - d$	$- d$	$-M - d$	$-4M - 100 - d$	$4M + 100 - d$	$-2M - d$

Note that we will not carry the zero requirement vector in the tableaus although its presence is to be understood. The simplex criterion for the perturbed problem dictates that k_5 replace l_3 as a basic variable. We perform the indicated operations on the rows of the initial tableau

to obtain the second tableau.

Second Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_4	k_5	k_6	k_7
				1	$-W$	$-W$	$-W$	d	d	d	d	d	d	d
k_0	1	0	100	1	0	0	100	0	100	0	0	0	0	-100
l_1	$-W$	1	0	0	1	0	-1	0	-2	-1	1	0	-1	3
l_2	$-W$	1	-1	0	0	1	-2	1	-3	1	0	0	1	3
k_5	d	0	1	0	0	0	1	-1	1	0	0	1	-1	-1
				0	0	0		$-W-2d$	$5W+100$	$-d$	$-W-d$	0	$-2d$	$-6W-100-2d$

Operations on rows of first tableau

$(0) = (0) + 100\ (3)$ $\qquad$ $(1) = (1) - (3)$

$(2) = (2) - 2\ (3)$ $\qquad$ $(3) = (3)$

$(4) = (4) + \left[4W + 100 + d \right] (3)$

We note that the last row of the columns corresponding to artificial variables is irrelevant for our purposes and may be discarded whenever convenient; we will do this for each artificial variable as it is removed from the basis. The simplex criterion for the perturbed problem dictates that k_7 replace l_2 as a basic variable. We perform the indicated transformation to obtain the third tableau.

Third Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_4	k_5	k_6	k_7
				1	$-W$	$-W$	$-W$	d	d	d	d	d	d	d
k_0	1	33.3	66.7	1	0	33.3	33.3	33.3	0	33.3	0	0	33.3	0
l_1	$-W$	0	1	0	1	-1	1	-1	1	2	1	0	-2	0
k_7	d	.33	-.33	0	0	.33	-.67	.33	-1	.33	0	0	.33	1
k_5	d	.33	.67	0	0	.33	.33	-.67	0	.33	0	1	-.67	0
				0	0			$W+(100-4d)/3$	$-W-2d$	$2W+(100-d)/3$	$-W-d$	0	$2W+(100-4d)/3$	0

Operations on rows of second tableau

(0) = (0) + 33.3 (3)
(1) = (1) - (2)
(2) = .33 (2)
(3) = (3) + .33 (2)
(4) = (4) + .33 $\left[\, 6W + 100 + 2d \,\right]$(2)

The simplex criterion for the perturbed problem dictates that k_2 replace l_1 as a basic variable. We perform the transformation and obtain the fourth tableau.

Fourth Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_4	k_5	k_6	k_7
				1	$-M$	$-M$	$-M$	d	d	d	d	d	d	d
k_0	1	33.3	66.7	1	0	33.3	33.3	33.3	0	33.3	0	0	33.3	0
k_2	d	0	1	0	1	-1	1	-1	1	-2	1	0	-2	0
k_7	d	.33	.67	0	1	-.67	.33	-.67	0	-1.67	1	0	-1.67	1
k_5	d	.33	.67	0	0	.33	.33	-.67	0	.33	0	1	-.67	0
				0				$10(10-d)/3$	0	$(100-10d)/3$	d	0	$4(25-4d)/3$	0

Operations on rows of third tableau

(0) = (0)　　　(1) = (1)

(2) = (2) + (1)　　　(3) = (3)

(4) = (4) + $\left[M + 2d \right]$ (1)

Since $y_{4\ 3+r} \geq 0$ for r = 1,...,7, this tableau represents an optimal solution and we have established that $AU + B > 0$ is consistent.

We now utilize the procedure of section III B to determine if the set satisfying $AU + B > 0$ is unbounded. We form the fifth tableau as indicated below.

Fifth Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_4	k_5	k_6	k_7
				1	-M	-M	-M	d	d	d	d	d	d	d
k_0	1	33.3	66.7	1	0	33.3	33.3	33.3	0	33.3	0	0	33.3	0
k_2	d	0	1	0	1	-1	1	-1	1	-2	1	0	-2	0
k_7	d	.33	.67	0	1	-.67	.33	-.67	0	-1.67	1	0	-1.67	1
k_5	d	.33	.67	0	0	.33	.33	-.67	0	.33	0	1	-.67	0
				0				-10d/3	0	-13d/3	0	0	-16d/3	0

Operations on rows of fourth tableau

(m) = (m) for m = 0,...,3

(4) = (4) - (1) (last seven columns only)

The heavy line under the first row of the tableau is to remind us that the first row is to be considered zero for purposes of determining unboundedness; it is carried in its present form only in the hope that it will provide an extreme point solution for the nonhomogeneous system while checking the consistency of the homogeneous system. No more computations are required here because already in this tableau we have indications of unbounded solutions to the new linear programming problem under k_1 and k_6 which signal that $AU > 0$ is inconsistent. Furthermore the fact that all

$N = 3$ basis vectors are required to specify the vectors corresponding to k_1 and k_6 signals that $AX \geq 0$ admits to only the trivial solution. Therefore the set satisfying $AX + B > 0$ is bounded.

VII C. Procedure for Section IV: We now may check the various constraints for redundancy. Returning to the fourth tableau we note that the constraint corresponding to k_4 satisfies the conditions of Theorem VI. It is therefore redundant and may be discarded. After discarding this constraint, we note that the constraints corresponding to k_2, k_5, and k_7 satisfy the conditions of Theorem VIII. Therefore none of them are redundant.

To check the constraint corresponding to k_3 for redundancy we begin applying the dual simplex algorithm to the fourth tableau (less the column of k_4) using the column of k_3 as the requirement vector. We select k_2 to be replaced by k_6. Performing the transformation yields the sixth tableau:

Sixth Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_3	k_5	k_6	k_7
				1	$-d$	$-d$	$-d$	d	d	d	d	d	d
k_0	1	33.3	83.3	1	16.7	16.7	50	16.7	16.7	0	0	0	0
k_6	d	0	-.5	0	-.5	.5	-.5	.5	-.5	1	0	1	0
k_7	d	.33	-.17	0	.167	.167	-.5	.167	-.33	0	0	0	1
k_5	d	.33	.33	0	-.33	.67	0	-.33	-.33	1	1	0	0
				0				$(50-2d)/3$	$(50-2d)/3$	v	0	0	0

Operations on the rows of the fourth tableau

(0) = (0) + 16.7(1)   (1) = -.5(1)

(2) = (2) - .33(1)   (3) = (3) - .33(1)

(4) = (4) + [16.7 - 2.67d] (1)

In this tableau we have the condition which by Theorem VII justifies our discarding the constraint corresponding to k_3 as redundant. With this column removed, we note that the constraints corresponding to k_5, k_7, and in particular k_6 satisfy the conditions of Theorem VIII. Therefore none of them are redundant.

To check the constraint corresponding to k_1 for redundancy we apply the dual simplex algorithm to sixth tableau (less the column of k_3) using the column of k_1 as a

requirement vector. We select k_5 to be replaced by k_2. The transformation results in the seventh tableau.

Seventh Tableau

				k_0	l_1	l_2	l_3	k_1	k_2	k_5	k_6	k_7
				1	-M	-M	-M	d	d	d	d	d
k_o	1	50	100	1	0	50	50	0	0	50	0	0
k_6	d	-.5	-1	0	0	-.5	-.5	1	0	-1.5	1	0
k_7	d	-.5	-1	0	1	-1.5	-.5	1	0	-2.5	0	1
k_2	d	-1	-1	0	1	-2	0	1	1	-3	0	0
				0				$\bar{d}$	o	[illegible]	o	o

Operations on rows of sixth tableau

(0) = (0) + 50 (3)　　　(1) = (1) - 1.5 (3)

(2) = (2) - 2.5 (3)　　　(3) = - 3 (3)

(4) = (4) + [50 - 3d] (3)

In this tableau we have a sufficient condition to discard the constraint corresponding to k_1 as redundant.

VII D. Procedure for Section V: We turn now to the procedure for generating extreme points, first constructing a table with n + 4 columns (plus 1 for the equality constraints) and $\binom{n}{m} = \binom{4}{3} = 4$ rows. We then enter the extreme point solutions already obtained which are real (v_c, v_e, v_f v_d, v_b).

Constraints						
k_2	k_5	k_6	k_7	v_a	v_b	tableau
			+			
33.3	33.3	+	0	33.3	66.7	fourth
50	+	50	0	50	100	seventh
+	50	16.7	16.7	33.3	33.3	sixth

We start the generation process with the fourth tableau and find

1. The adjacent extreme point found by replacing k_2 is one for which the constraints associated with k_5, k_6 and k_7 hold as equalities. This point was obtained in the sixth tableau.
2. The adjacent extreme point found by replacing k_5 is one for which the constraints associated with k_2, k_6, and k_7 hold as equalities. This point was obtained in the seventh tableau.
3. The adjacent extreme point found by replacing k_7 is one for which the constraints associated with k_2, k_5, and k_6 hold as equalities. This point has not yet been obtained.

To obtain this last point we replace k_7 by k_6. The 'value' information and the last row of the tableau are not needed for generating extreme points. Therefore we discard those

sections of the tableau. The transformation carried out on the fourth tableau less the columns of k_1, k_3, and k_4 produces the eighth tableau:

Eighth Tableau

			k_0	l_1	l_2	l_3	k_2	k_5	k_6	k_7	Operations on rows of the 4-th tableau
k_0	40	80	1	20	20	40	0	0	0	20	(0) = (0) + 20 (2)
k_2	-.4	.2	0	-.2	-.2	.6	1	0	0	-1.2	(1) = (1) - 1.2 (2)
k_6	-.2	-.4	0	-.6	-.4	-.2	0	0	1	-.6	(2) = -.6 (2)
k_5	.2	.4	0	-.4	.6	.2	0	1	0	-.4	(3) = (3) - .4 (2)

We may now enter this extreme point solution in the appropriate row of the table. Since this accounts for all the combinations of four constraints taken three at a time, it is not necessary to determine the adjacent extreme points for tableaus six, seven, and eight.

VII E. Procedure of Section VI: Suppose that we now wish to restrict the convex set more stringently by requiring

$$.107\ v_c - .016\ v_e - .169\ v_f - 3.58 \geq 0$$

Following the procedure of section VI we first form the vector

$$(B_3 \vdots A_3)' = (-3.58, -.169, -.016, .107)$$

We may incorporate this additional constraint into the seventh

tableau (the last one for which we have a y_{4r} row) multiplying $(B_3 \mid A_3)'$ by G' which is the 4x4 matrix in the columns of k_0, l_1, l_2, and l_3.

$$\begin{pmatrix} 1 & 0 & 50 & 50 \\ 0 & 0 & -.5 & -.5 \\ 0 & 1 & -1.5 & -.5 \\ 0 & 1 & -2 & 0 \end{pmatrix} \begin{pmatrix} -3.58 \\ -.169 \\ -.016 \\ .107 \end{pmatrix} = \begin{pmatrix} .960 \\ -.045 \\ -.198 \\ -.138 \end{pmatrix}$$

Assigning k_8 to the new constraint we now add a new column to the seventh tableau (less the column of k_1) to obtain the ninth tableau.

Ninth Tableau

				k_0	l_1	l_2	l_3	k_2	k_5	k_6	k_7	k_8
				1	$-M$	$-M$	$-M$	d	d	d	d	d
k_0	1	50	100	1	0	50	50	0	50	0	0	.96
k_6	d	-.5	-1	0	0	-.5	-.5	0	-1.5	1	0	-.045
k_7	d	-.5	-1	0	1	-1.5	-.5	0	-2.5	0	1	-.198
k_2	d	-1	-1	0	1	-2	0	1	-3	0	0	-.138
				0				0	$50-3M$	0	0	$.96-1.33M$

After computing the last entry in the column of k_8, we can conclude that the new constraint has not introduced an inconsistency because the linear programming optimality

conditions are satisfied. The new convex set must be bounded because the original set was. We still do not know if the new constraint has actually restricted the convex set more stringently; it may be redundant. On the other hand its inclusion may have rendered some of the original constraints redundant. Clearly the constraints corresponding to k_2, k_6, and k_7 are not redundant because they satisfy the conditions of Theorem VIII in the ninth tableau. Therefore only the constraint corresponding to k_5 could have been made redundant by the addition of the new constraint. We may check this possibility by applying the dual simplex algorithm to the ninth tableau using the column of k_5 as the requirement vector. We select k_2 to be replaced by k_8 and perform the transformation to obtain the tenth tableau.

Tenth Tableau

				k_0	l_1	l_2	l_3	k_2	k_5	k_6	k_7	k_8
				1	-M	-M	-M	d	d	d	d	d
k_0	1	43.0	93.0	1	6.95	36.1	50	6.95	29.1	0	0	0
k_6	d	-.171	-.671	0	-.329	-.158	-.5	-.329	-.514	1	0	0
k_7	d	.936	.436	0	-.436	1.77	-.5	-1.44	1.31	0	1	0
k_2	d	7.24	7.24	0	-7.24	14.5	0	-7.24	21.7	0	0	1
				0				6.95-10d	29.1-22d	0	0	0

Operations on rows of ninth tableau

(0) = (0) + 6.95 (3)

(1) = (1) -.329 (3)

(2) = (2) - 1.44 (3)

(3) = -7.24 (3)

(4) = (4) + [6.95 - 10d] (3)

We have yet to establish the redundancy or nonredundancy of the constraint corresponding to k_5 in this tableau. Another use of the dual simplex algorithm produces the eleventh tableau for which k_2 has replaced k_6 as a basic variable.

Eleventh Tableau

				k_0	l_1	l_2	l_3	k_2	k_5	k_6	k_7	k_8
				1	-d	-d	-d	d	d	d	d	d
k_0	1	39.5	78.9	1	0	39.5	39.5	0	18.3	21.1	0	0
k_2	d	.521	2.04	0	1	-.479	1.52	1	1.56	-3.04	0	0
k_7	d	1.68	3.37	0	1	.684	1.68	0	4.05	-4.37	1	0
k_8	d	11.0	22.0	0	0	11.0	11.0	0	53.0	-22.0	0	1
				0				0	18.3-37.6d	21.1-30.4d	0	0

Operations on rows of tenth tableau

(0) = (0) + 21.1 (1)   (1) = - .304 (1)

(2) = (2) - 4.37 (1)   (3) = (3) - 22.0 (1)

(4) = (4) + $\left[\, 21.1 - 30.4\, d \,\right]$ (2)

In this tableau we have established that the constraint corresponding to k_5 is redundant while the new constraint is not redundant.

Because we are left with four constraints the problem of specifying extreme points is the same as the one encountered

earlier. We form a table and enter the extreme point solution already obtained. These are read in the table (v_c, v_e, v_f v_d, v_b).

Constraints						
k_2	k_6	k_7	k_8	v_d	v_b	Tableau
50	50	0	+	50	100	Ninth
		+				
39.5	+	39.5	0	39.5	79.5	Eleventh
+	50	36.1	6.95	40.0	93.0	Tenth

We have only one extreme point unaccounted for, and it may be obtained from the ninth tableau (less the column of k_5) by replacing k_7 with k_8 as a basic variable. This replacement leads to the twelfth tableau.

Twelfth Tableau

			k_o	l_1	l_2	l_3	k_2	k_6	k_7	k_8
k_o	47.6	95.2	1	4.84	42.8	47.6	0	0	4.34	0
k_6	-.386	-.771	0	-.229	-.157	-.386	0	1	-.229	0
k_8	2.52	5.04	0	-5.04	7.56	2.52	0	0	-5.04	1
k_2	-.652	-.304	0	.303	-.955	.348	1	0	-.696	0

Operations on rows of ninth tableau

(0) = (0) + 4.84 (2)

(1) = (1) - .229 (2)

(2) = - 5.04 (2)

(3) = (3) - .696 (2)

We may now complete our list of the extreme points of this convex set by entering the extreme point solution from the twelfth tableau in the second row of the table.

Bibliography

Ackoff, R. L., *Scientific Method*, N. Y., Wiley, 1962

Birnbaum, A. and Maxwell, A. E., "Classification procedures based on Bayes's formula," *Applied Statistics*: 9 (1960) 152-169

Carver, W. B., "Systems of Linear Inequalities," *Annals of Mathematics*, 23 (1921-22), 212-220

Chernoff, H. and Moses, L. E., *Elementary Decision Theory*, N. Y., Wiley, 1959

Chow, C. K., "An Optimal Character Recognition System Using Decision Functions," I. R. E. *Transactions on Electronic Computers*, Vol. EC-6, No. 4 (Dec. 1957), 247.

Churchman, C. W., *Prediction and Optimal Decision*, Englewood Cliffs, N. J., Prentice-Hall, 1961

Churchman, C. W., Ackoff, R. L. and Arnoff, E. L., *Introduction to Operations Research*, N. Y., Wiley, 1957

Flagle, C. D. and Lechat M., "Statistical Decision Theory and the Selection of Diagnostic and Therapeutic Strategies in Public Health," *Proceedings* of the Third Meeting of the International Federation of Operations Research Societies, Oslo, July 1963

Graybill, F. A., *An Introduction to Linear Statistical Models*, Vol. 1, N. Y., McGraw-Hill, 1961

Hadley, G., *Linear Programming*, Reading, Mass., Addison-Wesley, 1962

Luce, R. D. and Raiffa, H., *Games and Decisions*, N. Y., Wiley, 1957

Overall, J. E. and Williams, C. M., "Conditional Probability Model for Diagnosis of Thyroid Function," *Journal of the American Medical Association*, Vol. 183 (1963), 307-313

Raiffa, H. and Schlaifer, R., *Applied Statistical Decision Theory*, Boston, Division of Research, Graduate School of Business Administration, Harvard University, 1961

Stillson, P., "A Method for Defect Evaluation," *Industrial Quality Control*, XI, No. 1 (July, 1954), 9-12

Vita

The author was born in Lewisburg, Pennsylvania on 17 October 1937. He attended the University of Nebraska (1954-55) and the United States Naval Academy (1955-59). In 1960 he was assigned to The Johns Hopkins University under the Junior Line Officer Advanced Scientific Educational (Burke) Program of the United States Navy.

thesB736
The sensitivity of optimal strategies in
3 2768 001 01682 7
DUDLEY KNOX LIBRARY